Immunoglobulin light chains in Systemic Lupus Erythematosus

Fraser, Louise

Awarding institution:
King's College London

Immunoglobulin light chains in Systemic Lupus Erythematosus

By

Louise Fraser

Supervised by Professor Jo Spencer and

Professor Andrew Cope

Peter Gorer Department of Immunobiology, School of Medicine

King's College London, London, UK.

A thesis submitted in fulfillment of the requirements for the Degree of Doctor of Philosophy in Immunology at King's College London, University of London

August 2013

Declaration

I declare that the work described in this thesis has been personally prepared and all experiments have been personally carried out unless otherwise cited. Sources of information used during the production of this thesis are acknowledged by means of a reference.

Louise Fraser

August 2013

Acknowledgements

I would firstly like to thank my fantastic supervisor Professor Jo Spencer for giving me the opportunity to study in her lab under an Oliver Bird funded PhD Studentship. Jo not only offered her guidance, support, patience and encouragement with this project on a daily basis, but taught me to start having confidence in my own scientific ability. I would also like to thank Professor Andrew Cope, who too offered continuous support and welcomed meetings to discuss the project and new potential experiments.

I greatly appreciated the regular meetings with my thesis committee members; Professor David D'Cruz and Dr. Michael Robson, who offered guidance and suggestions to ensure my thesis progression ran efficiently.

I would also like to greatly thank Dr. Richard Ellis and Thomas Hayday for their invaluable help with flow cytometry and cell sorting throughout this project. My thanks also extend to Lee Meng Choong for facilitating the collection of SLE blood samples from patients attending the Louise Coote Lupus Unit at St. Thomas' Hospital, London.

My warm thanks also go to members and friends of the Spencer Lab; Dr. Yuan Zhao, Dr. Jessica Thomas, Dr. Pamela Lutalo, Romeeza Tahir, Dr. Manu Shankar-Hari and Dr. Dia Chavele. I would also like to thank Saba Nayar and Dr. Francesca Barone who supported me experimentally at the start of this project. My time at King's has been made unforgettable thanks to my lovely friends of the department who have made time away from the lab full of laughter and kindness.

Finally, I would like to thank the Oliver Bird Rheumatism Programme for funding this work and providing an additional source of support and encouragement at the annual Oliver Bird Conferences which were on every occasion, stimulating and enjoyable.

Abstract

Systemic Lupus Erythematosus (SLE) is a chronic autoimmune disease of elusive origin and characterised by polyclonal B cell hyperactivity and the production of pathogenic antibodies targeting self DNA and nucleoproteins. Clinical manifestations of SLE are highly heterogeneous and include multisystem inflammation of connective tissue and vasculitis of the central nervous system (CNS). SLE is known to have genetic associations, and is widely acknowledged to involve a profound breakdown in immune tolerance. The aim of this project was to identify whether defects in early B cell development that affect the expressed repertoire of immunoglobulin light chains could be observed and to ask whether receptor editing contributes to disease pathogenesis.

Chapter 3 describes a high-throughput sequencing analysis of the human immunoglobulin kappa light chain gene repertoire in healthy and SLE mature naive peripheral B cells. We observed that involvement of gene segments as the repertoire develops is equivalent in health and SLE and that a previously described bias towards usage of the gene segment *IGKV4-1* in SLE is only observed in the expressed repertoire analysed.

Chapter 4 describes an inefficient function of the kappa deleting element (KDE) in SLE, which manifests as reduced frequency of KDE rearrangement status in populations of $CD19^{+}$ B cells and a failure to inactivate alleles of *IGK* allowing them to accumulate somatic hypermutations within non-productive *IGK* rearrangements.

Chapter 5 identifies a potential biological outcome of a failure to inactivate rearranged alleles of *IGK* in SLE; cell surface expression of both kappa and lambda light chains were observed by flow cytometry analyses. Efforts to disprove this were unsuccessful and

detection of both light chain transcripts was confirmed through single cell PCR amplification of cDNA.

The data in this thesis suggest that the failure of the KDE to inactivate alleles of *IGK* efficiently in SLE permits the expression of *IGK* light chains that have been selected against resulting in allelic inclusion. This mechanism may account for the broader and less stringently regulated antigen-binding B cell repertoire associated with SLE.

Table of Contents

List of Figures

Figure 5.10 Flow cytometry analysis of the frequency of dual light chain-expressing B cells in different B cell subsets.

Figure 5.11 The correlation between rearrangement status of the KDE and cell surface light chain expression of both Igκ and Igλ light chains by $CD19^+$ B cells in SLE.

Figure 5.12 The correlation between the frequency of dual light chain expression by $CD19^+$ SLE B cells and the concentration of free serum light chains.

List of Tables

List of acronyms

(C)CIA	(Chronic) collagen induced arthritis
A nucleotide	Adenosine
α4β7	Alpha4 beta7 integrin
aa	Amino acid
ABCB1	ATP-binding cassette, sub-family B (MDR/TAP), member 1
Abs	Antibodies
Ags	Antigens
AID	Activation induced cytidine deaminase
ANA	Anti-nuclear antibody
APC	Allophycocyanin
APC	Antigen-presenting cell
APE	Apurinic/apyrimidinic endonuclease
APRIL	A proliferation-inducing ligand
ASC	Antibody secreting cell
ATM	Ataxia telangiectasia mutated
AZA	Azathioprine
BAFF	B-cell activating factor
BAFF-R	B-cell activating factor receptor
BANK1	B-cell scaffold protein with ankyrin repeats 1.
Bcl-6	B-cell lymphoma 6
BCMA	B cell maturation ligand
BCR	B-cell receptor
BILAG	British Isles Lupus Assessment Group index
Blimp-1	B lymphocyte-induced maturation protein-1
BLK	B-lymphoid tyrosine kinase
BLyS	B lymphocyte stimulator
BM	Bone marrow
bp	Base pairs
Btk	Bruton's tyrosine kinase
C nucleotide	Cytosine
C region	Constant region
CCL19	Chemokine (C-C motif) ligand 19
CCL21	Chemokine (C-C motif) ligand 21
CCR7	Chemokine (C-C motif) receptor 7
CD10	Cluster of differentiation 10
CD138	Cluster of differentiation 138, or syndecan-1
CD14	Cluster of differentiation 14
CD19	Cluster of differentiation 19
CD20	Cluster of differentiation 20
CD21	Cluster of differentiation 21
CD23	Cluster of differentiation 23
CD24	Cluster of differentiation 24
CD25	Cluster of differentiation 25
CD27	Cluster of differentiation 27
CD3	Cluster of differentiation 3
CD38	Cluster of differentiation 38
CD4	Cluster of differentiation 4
CD40	Cluster of differentiation 40
CD40L	Cluster of differentiation 40 ligand, or CD154

CD8	Cluster of differentiation 8
CD86	Cluster of differentiation 86
cDNA	Complementary deoxyribonucleic acid
CDR	Complementary determining region
CFA	Complete Freunds' adjuvant
CSR	Class switch recombination
CT	Cycle threshold (qRT-PCR)
CXCL13	C-X-C motif chemokine 13
CXCR5	C-X-C motif chemokine 13
CYC	Cyclophosphamide
DC	Dendritic cell
DN	Double negative
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
dsDNA	Double stranded DNA
DTH	Delayed type hypersensitivity
EAE	Experimental autoimmune encephalomyelitis
EBV	Epstein-Barr virus
ELISA	Enzyme-linked immunofluorescence assay
ERK/MAP kinase	Extracellular signal-regulated kinase/mitogen activated protein kinases
Fab	Fragment antibody binding
FACS	Fluorescence activated cell sorting
Fc	Fragment crystallisable
FCS	Fetal calf serum
FDC	Follicular dendritic cell
FITC	Fluorescein isothiocyanate
FL	Follicular lymphoma
FMO	Fluorescence minus one
FO B	Follicular B cell
FoxP3	Forkhead box P3
FSC	Forward scatter
FW	Framework region
G nucleotide	Guanine
GALT	Gut-associated lymphoid tissue
GC	Germinal centre
GFP	Green fluorescence protein
GPI	Glucose-6-phosphate isomerase
GWAS	Genome wide association studies
HCQ	Hydroxychloroquine
HEV	High endothelial venule
HIGM2	Hyper-IgM syndrome
HMPB1/2	High-mobility group proteins B1 and B2
HSC	Hemopoetic stem cell
HTS	High-throughput sequencing
IC	Immune complex
IFA	Immunofluorescence assay
IFN-α	Interferon alpha
IFN-β	Interferon beta
IFN-γ	Interferon gamma
Ig	Immunoglobulin
Ig-α/Ig-β	CD79a/CD79b
IGH	Immunoglobulin heavy chain

IGHD	Immunoglobulin heavy chain diversity region
IGHJ	Immunoglobulin heavy chain joining region
IGHV	Immunoglobulin heavy chain variable region
IGK	Immunoglobulin kappa light chain
IGL	Immunoglobulin lambda light chain
IL-10	Interleukin-10
IL-7	Interleukin-7
IL-7R	Interleukin-7 receptor
IRAK1	Interleukin-1 receptor-associated kinase
IRF5	Interferon regulatory factor 5
iRSS	Intronic recombination signal sequence
iRSS-KDE	Intronic recombination signal - KDE rearrangement
iTREG	Inducible regulatory T cell
KDE	Kappa deleting element
LN	Lymph node
LPS	Lipopolysaccharride
LYP	Lymphoid tyrosine phosphatase
MACS	Magnetic cell separation
MID	Multiplex identifier
MMF	Mycophenolate mofetil
mRNA	Messenger RNA
MSH2/MSH6	DNA mismatch repair proteins 2 and 6
MTX	Methotrexate
MZ B	Marginal zone B cell
NET	Neutrophil extraellular trap
Nrp-1	Neuropilin-1
nTREG	Natural regulatory T cell
NZB/NZW	New Zealand black/ New Zealand white
PAMP	Pathogen-associated molecular pattern
PBMC	Peripheral blood mononuclear cell
PC	Plasma cell
PCR	Polymerase chain reaction
pDC	Plasmacytoid dendritic cell
PE	Phycoerythrin
PerCP-Cy5.5	Peridinin-chlorophyll-protein complex (PerCP) conjugated with cyanine dye 5.5
Pred	Prednisolone
PRR	Pattern recognition receptor
PTPN22	Protein tyrosine phosphatase, non-receptor 22
qRT-PCR	Quantitative real-time polymerase chain reaction
RA	Rheumatoid arthritis
RAG 1/2	Recombination activating genes 1 and 2
RNA	Ribonucleic acid
RNA-FISH	RNA-fluorescent in situ hybridisation
RNP	Ribonucleoprotein
ROR-γt	RAR-related orphan receptor gamma t
RSS	Recombination signal sequence
RTX	Rituximab
RUNX1	Runt-related transcription factor 1
S region	Switch region
SHM	Somatic hypermutation
SLE	Systemic lupus erythematosus
SLEDAI	Systemic Lupus Erythematosus Disease Activity Index

SLyRT	Sort-Lysis Reverse Transcription buffer
Sm	Smith antigen
SNP	Single nucleotide polymorphism
ssDNA	Single stranded DNA
STAT	Signal transducer and activator of transduction
Syk	Spleen tyrosine kinase
T nucleotide	Thymidine
T1 B	Transitional 1 B cell
T1D	Type 1 diabetes
T2 B	Transitional 2 B cell
T3 B	Transitional 3 B cell
TACI	Transmembrane activator-1 and calcium modulator and cyclophilin ligand-interactor
TCR	T cell receptor
TD	Thymus-dependent
TdT	Terminal deoxynucleotidyl transferase
TFH	T follicular helper cells
Tg	Transgenic
TGFβ	Tumour necrosis factor beta
Th1	Type 1 T helper cells
Th17	Type 17 T helper cells
Th2	Type 2 T helper cells
TI	Thymus-independent
TLR	Toll-like receptor
TNFα	Tumour necrosis factor alpha
TNFAIP3	Tumor necrosis factor, alpha-induced protein 3
TNFR	Tumour necrosis factor receptor
U nucleotide	Uracil
UNG	Uracil-DNA glycosylase
V(D)J	Variable (diversity) joining recombination
WT	Wild-type

Chapter 1
Introduction

1. The human immune system

A central feature of the mammalian immune system is the ability to recognise and respond specifically to foreign antigen whilst maintaining tolerance to self. The innate immune system achieves this by expression of pattern recognition receptors (PRRs) that recognise pathogen-associated molecular patterns (PAMPs) not shared by mammalian cells. In contrast, specific antigen recognition in the adaptive immune system is mediated by cell-unique antigen receptors expressed by T cells and B cells. The mechanisms generating the antigen receptors of the adaptive immune response include some random genetic events that maximise the potential for receptor diversity. However, unlike the innate immune system that is focused on exogenously derived antigens, the receptors of the adaptive immune system comprise a diverse repertoire of cell surface expressed T cell receptors (TCR) and immunoglobulin (Ig), some of which have the potential to recognise self. The various mechanisms that regulate recognition of self by the adaptive immune system are therefore important and failure can lead to disease. Experiments in this thesis investigate regulation of expression of immunoglobulin light chains by human B cells and how this system might fail in the autoimmune disease systemic lupus erythematosus (SLE).

The introduction to this thesis will be separated into two sections; the first will describe how the nascent human B cell pool is established and the different fates met by B cells during different stages of their development. The mechanisms that serve to regulate human B cell development, maintain monoclonal expression of antigen receptors and counter-select newly produced B cells expressing unwanted specificities will also be discussed. The second section will describe the role of B cells in human biological states where self-tolerance fails with particular focus on SLE.

1.1 The development of a human immunoglobulin

Developing B cell subsets are defined by the sequential stages of immunoglobulin gene rearrangement. Therefore, the description of B cells will start by analysis of the immunoglobulin loci.

1.1.1 The human immunoglobulin loci

Tonegawa's pioneering studies of the 1980s described that the Ig genes were encoded by discrete, non-functional gene segments that assemble by somatic recombination to form functional receptors capable of recognising cognate antigen, known as V(D)J recombination (Tonegawa, 1983). This in the first instance accounts for the extreme repertoire diversity required by Ig B cell repertoires for a robust immune response. The human Ig gene loci are located on different chromosomes, each possessing unique organisation. The human Ig heavy chain (*IGH*) gene locus is encoded on chromosome 14q32.33 over four regions, where the variable (V_H), diversity (D_H), joining (J_H) and constant (C_H) gene segments are located in sequential clusters (Lefranc, 2001a). In the human, there are approximately 40 V_H segments, 25 D_H segments, 6 J_H regions and a large cluster of 10 C_H segments, though only nine of the latter segments are functional due to the presence of a C_H pseudogene (Murphy. K, 2008). This differs from mice, where the *IGH* locus consists of approximately 150 V_H segments, 12 D_H segments, 4 J_H segments and 8 functional C_H regions (Corcoran, 2005). The genetic organisation of the human *IGH* locus is illustrated below (Figure 1.1)

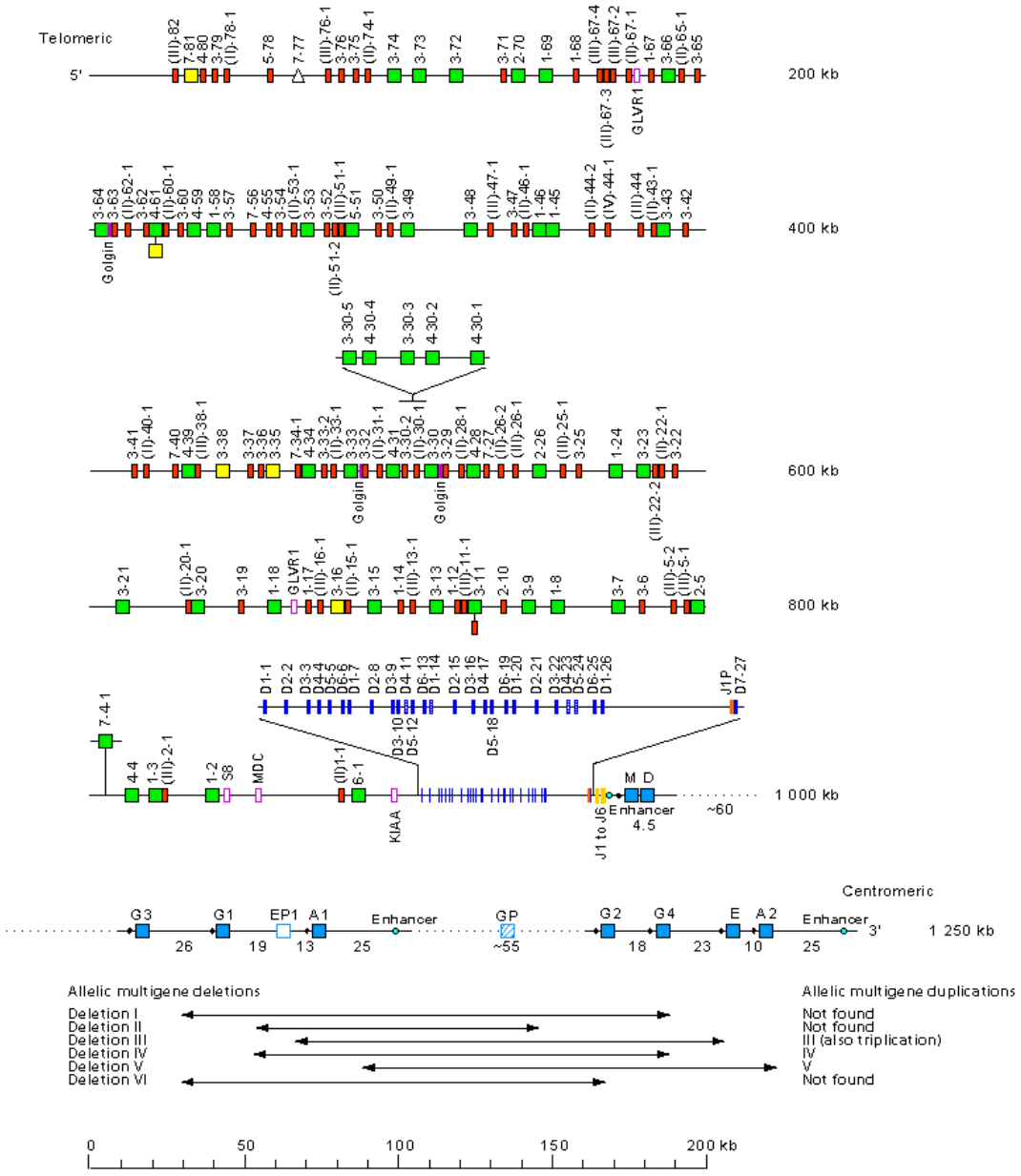


Figure 1.1 Locus representation of the human *IGH* locus. Figure reference: IMGT, the international ImMunoGeneTics information system.

The genes encoding the human kappa (*IGK*) and lambda (*IGL*) light chain loci, that will be referred to as Igκ and Igλ when describing the protein Ig molecule, are located on chromosomes 2p11.2 and 22q11.2 respectively, and differ from the *IGH* genes in that they do not possess D gene segments. The genetic organisation of the two light chain gene loci

also differs; *IGK* gene segments are ordered similarly to the heavy chain genes, V_κ-J_κ-C_κ where all J_κ segments lie 5' of the one C_κ region (Lefranc, 2001b). The human *IGK* locus consists of 40 functional V_κ gene segments, 5 J_κ segments and a single C_κ region (Murphy. K, 2008). The mouse *IGK* locus consists of ~140 V_κ gene segments, 5 J_κ segments and one C_κ gene segment (Corcoran, 2005). The human *IGL* locus is comprised of 30 functional V_λ gene segments, and 4 pairs of functional J_λ gene segments and C_λ genes (Murphy. K, 2008). Gene clusters at the *IGL* loci have a nested arrangement, where a cluster of V_λ gene segments is followed by J_λ segments which are associated with their own C_λ segment (Lefranc, 2001c). The genetic organisation of the human *IGK* and *IGL* loci are depicted below in Figure 1.2 and 1.3 respectively.

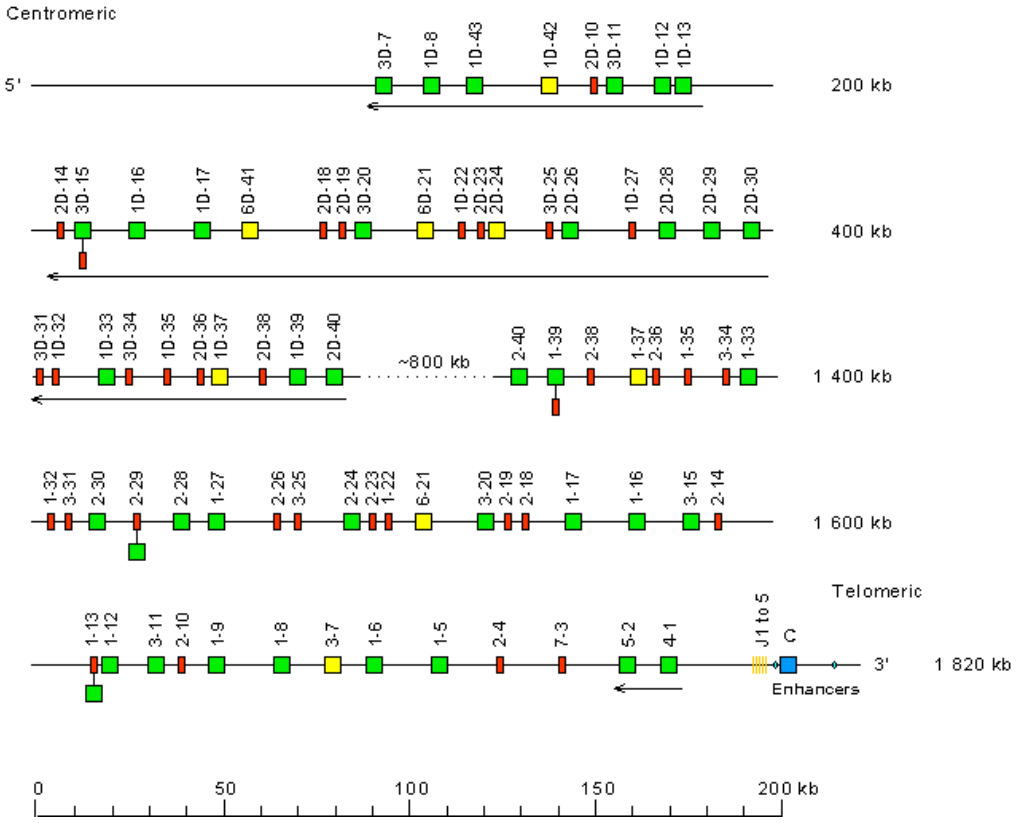


Figure 1.2 Locus representation of the human *IGK* light chain locus. Figure reference: IMGT, the international ImMunoGeneTics information system.

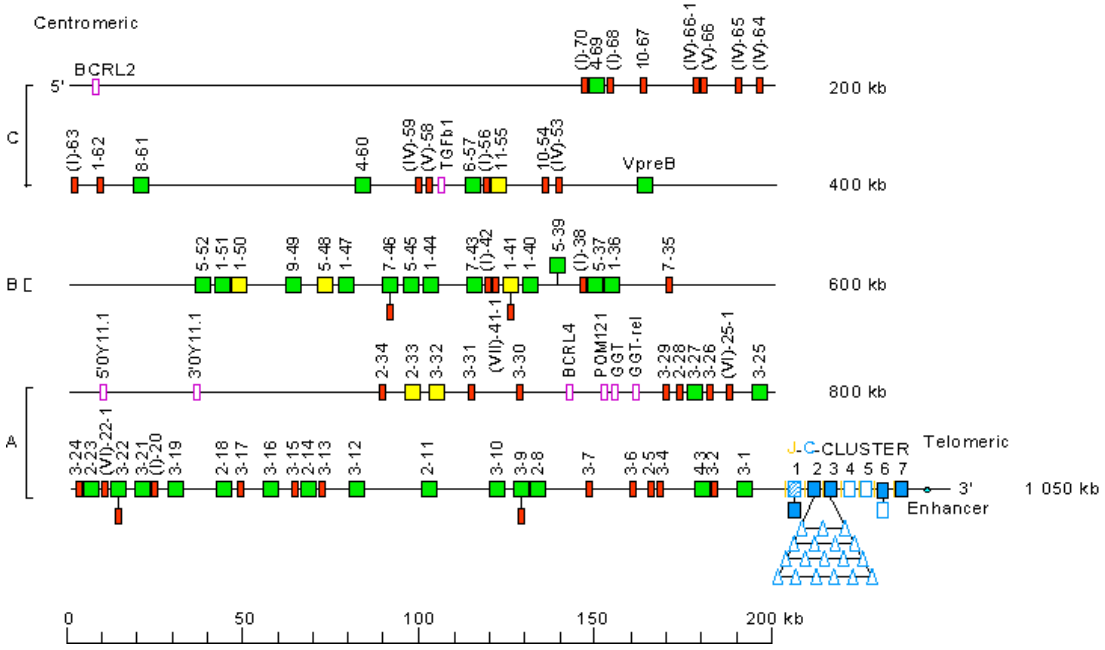


Figure 1.3 Locus representation of the human *IGL* light chain locus. Figure reference: IMGT, the international ImMunoGeneTics information system

1.1.2 Recombination signal sequences

Each functional V, D and J gene segment at both the heavy and light chain loci is associated with a recombination signal sequence (RSS), which is used by the cell's recombination machinery to direct the rearrangement process. RSSs consist of a heptamer sequence (typically CACAGTG) and nonamer sequence (typically ACAAAAACC) separated by a less conserved spacer sequence of either 12 or 23 nucleotides, considered to reflect one or two turns of the DNA double helix (Sakano et al., 1979; Schatz and Ji, 2011). The 12/23 rule is a mechanism that controls rearrangement, since a gene segment flanked by a 12 nucleotide spacer may only recombine to an RSS with a 23 nucleotide spacer, and vice versa (Eastman et al., 1996). Following successful recombination of a heavy and light chain, the immunoglobulin is assembled by linkage of two heavy chains (50-70 kDa) with

two light chains (approximately 23 kDa) by inter-chain disulphide bonds and other non-covalent interactions. Intra-disulphide bonds and similar non-covalent associations are present within each heavy and light chain, maintaining the structure of the Ig (Murphy. K, 2008). A schematic outlining the production of an Ig molecule is depicted in Figure 1.4. The enzymes and mechanisms that catalyse and regulate V(D)J recombination will be described in the section detailing B cell development.

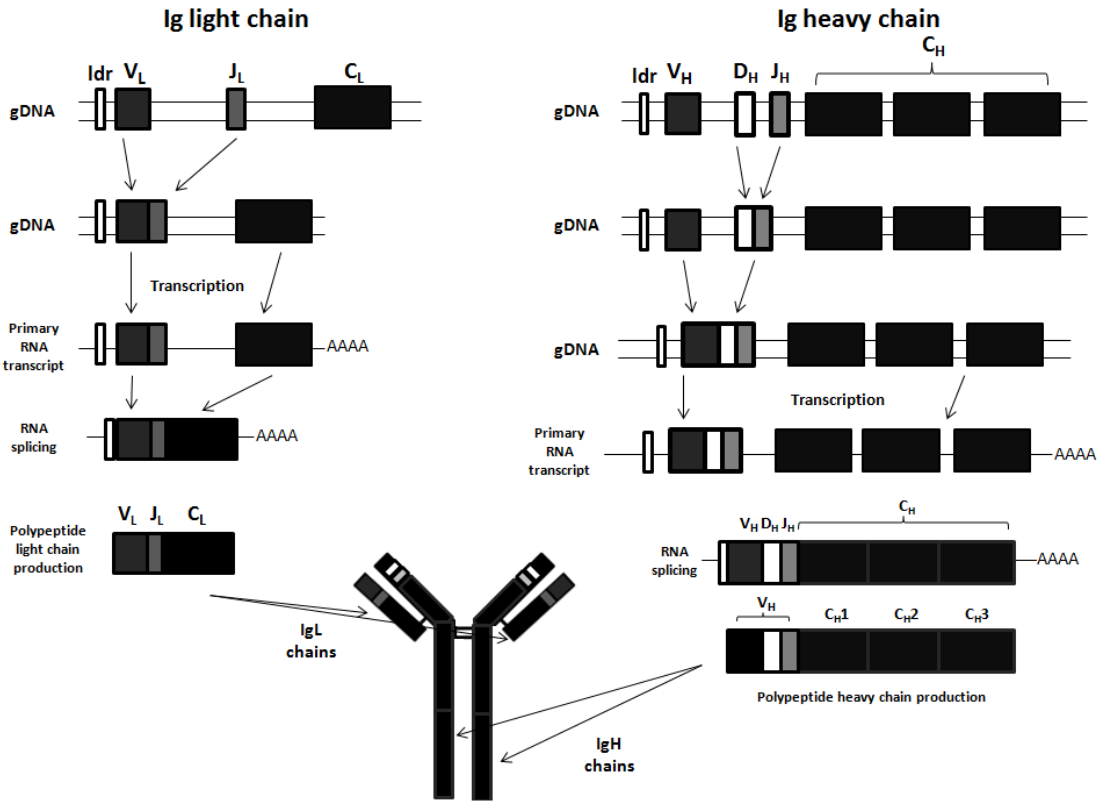


Figure 1.4 Immunoglobulin gene rearrangement at the heavy and light chain gene loci. Somatic recombination of the gDNA encoding the Ig chains takes place initially between the D_H and J_H gene segments at the heavy chain loci and the V_L and J_L segments at the light chain loci. D_HJ_H rearrangement takes place on both heavy chain alleles simultaneously, whereas V_LJ_L recombination takes place asynchronously at the light chain loci. The first allele on which D_HJ_H segments have rearranged then recombines with a V_H gene segment to complete rearrangement of the *IGH* variable region. V gene segments at both the heavy and light chain loci are preceded by a leader (ldr) sequence which is responsible for directing the newly rearranged polypeptide into the cells secretory pathways but is later cleaved following translation of the complete polypeptide. To complete rearrangement at the heavy chain locus C_H regions are spliced to the $V_HD_HJ_H$ region. Completion of light chain gene rearrangement involves the splicing of RNA to remove the ldr-V_L and J_L-C_L intronic regions for rearrangement of C_L to the V_LJ_L region.

1.1.3 The structure of human immunoglobulin

The antigen binding portion of an immunoglobulin is known as the variable (V) region and is located in the *Fab* fragment (fragment antigen binding) of the Ig molecule. It is comprised of the rearranged $V_HD_HJ_H$ segments of the heavy chain, known as the variable (V_H) region, and rearranged V_LJ_L of the either the *IGK* and *IGL* light chain loci, referred to as V_L region (Murphy. K, 2008). Variable regions are unique to each immunoglobulin, conferring individual specificity to each molecule predominantly through variations in their complementarity determining regions (CDRs); each immunoglobulin carries three CDRs on each heavy and light chain. CDR3 regions are separated by more conserved stretches of DNA known as framework regions (FW) which play a role in stabilisation of antigen binding.

The constant (C) regions of the heavy chains make up the *Fc* (Fragment crystallisable) portion of the Ig molecule, and are formed of discrete domains, C_H2, C_H3 and in some cases a C_H4 domain (C_H1 is a component domain of the *Fab* fragment) where conserved sequences within these domains determine the isotype of the immunoglobulin. The isotype ultimately determines the effector function of the secreted Ig. The genetic organisation of the *IGH* constant regions generally relates to the expression of Ig classes during B cell development. The C regions of immunoglobulin molecules are known to be the predominant sites that harbour N-glycosylated oligosaccharides (Dunn-Walters et al., 2000). The consensus sequence for N-glycosylation sites typically involve a Asn-X-Ser/Thr/Cys motif, where X may reflect any amino acid except a proline (Medzihradszky, 2008). A role of glycosylation in the maintenance of Ig structure and effector function was elucidated by a study involving removal of an attachment residue, Asn-297, located between the C_H2 domain of chimeric mouse-human IgG (Tao and Morrison, 1989). This study demonstrated that although the Ig lacking Asn-297 are assembled and secreted

normally, they are more sensitive to protease degradation than wild-type (WT) Ig and cannot bind FcγRI. It has also been identified that somatic hypermutation can affect N-glycosylation sites in *IGVH* genes used by mucosal plasma cells secreting IgM, IgA and IgG, where increased frequencies of hypermutation were found to be associated with the loss of germline encoded glycosylation sites (Dunn-Walters et al., 2000). Variable Ig regions have also been found to undergo glycosylation and is a subject of great interest due to the influence these modifications may have on antigen selection. Potential novel sites of N-glycosylation located at V_H CDR3 regions of B cells were identified in 79% of patients with follicular lymphoma (FL) compared to 9% in healthy controls following analysis of similar V_H CDR3 regions in control memory B and plasma cells carrying similar levels of mutation. This glycosylation was found to be introduced following somatic mutation in the germinal centre (GC), and it was proposed that the addition of carbohydrate at these sites supported tumour development by influencing antigen selection (Zhu et al., 2002).

A functional cell surface B cell receptor (BCR) exists in a complex through non-covalent association with the transmembrane Ig-α/Ig-β heterodimer, also referred to as CD79a/CD79b complex. This association is critical for mediating signal transduction to intracellular signalling molecules downstream of the BCR since the cytoplasmic tails of each C_H region of membrane-bound Ig molecules are short; mIgM and mIgD molecules contain only three amino acid residues, and the cytoplasmic tails of mIgA, mIgM and mIgE are 14, 28 and 28 residues respectively (Kindt, 2007; Reth, 1992). The Igα/Ig-β molecules have long cytoplasmic tails of 61 and 48 amino acids respectively and contain immunoreceptor tyrosine-based activation motifs (ITAMs) that contain two repeats of the consensus sequence Tyr-X-X-Leu/Ile (Isakov, 1997). BCR engagement by antigen induces rapid phosphorylation of tyrosine residues within the ITAMs that consequently provide docking sites for cytoplasmic effector molecules, predominantly Src family kinases, which

mediate the downstream signalling pathways that ultimately shape the gene expression profile of the activated B cell.

1.1.4 V(D)J recombination

The initiation of V(D)J recombination in the developing B cell is dependent on the introduction of chromosomal double-stranded DNA (dsDNA) breaks at RSS junctions of each V, D and J gene segment; an inherently harmful genetic reorganisation that is tightly controlled in a temporal, cell-lineage specific and stage-specific manner (Corcoran, 2005). V(D)J recombination is mediated by the protein products of the recombination activating genes (*RAG*s) which together are referred to as the V(D)J recombinase. Expression of V(D)J recombinase is restricted to cells of the lymphocyte lineage at specific stages in their development and associate with several other DNA binding a repair proteins, known collectively as high-mobility group proteins B1 and B2 (HMPB1/2) (Oettinger et al., 1990; Schatz and Ji, 2011).

The products of the *RAG* genes, RAG1 and RAG2, are nuclear proteins that mediate site-specific recombination of Ig gene segments preferentially between 12RSS and 23RSS signals. RAG1 and RAG2 have differing roles in Ig gene recombination. RAG1 has been demonstrated to activate V(D)J recombination in NIH 3T3 fibroblasts *in vitro* (Schatz et al., 1989), and co-transfection of *RAG1* and *RAG2* expression constructs was shown to be sufficient to activate recombination in many non-lymphoid cell types (Stanhope-Baker et al., 1996). The recognition of RSSs is considered to be mediated predominantly by RAG1, which carries an RSS nonamer-binding domain which anchors the RAG proteins to the DNA and an active site for the breakage of DNA. RAG1 is expressed throughout the cell cycle (Schatz and Ji, 2011). The central region of RAG1 interacts with the heptamer

sequence of the RSS and RAG2, and co-transfection of RAG2 into a fibroblast cell line already expressing RAG1 has been shown to confer a 1000-fold increase in the frequency of V(D)J recombination events (Oettinger et al., 1990). Thus, RAG2 is thought to enhance DNA cleavage by increasing the affinity and specificity of the RAG complex for RSSs, though it has negligible binding activity alone and is only expressed at stages G0 and G1 of the cell cycle. The C-terminus of RAG2 is also known to interact with histone proteins, where the introduction of mutations in RAG2 was demonstrated to severely inhibit the efficiency of V_H-D_HJ_H joining (West et al., 2005). Additional studies suggested that the RAG1/RAG2 complex in addition to other DNA binding proteins binds one RSS initially and 'captures' an unbound, complementary (dissimilar) RSS forming a synapse during recombination (Jones and Gellert, 2002).

Double stranded DNA cleavage involves sequential single strand nicking between the heptamer and Ig coding gene segment by hydrolysis of a phosphodiester bond of the DNA backbone (Stanhope-Baker et al., 1996). This generates a free 3' hydroxyl (OH^-) group which then attacks the remaining un-nicked DNA strand forming a hairpin loop. Two broken-ended DNA intermediates are produced; a signal end, blunt and 5' phosphorylated, and a coding end which is sealed in a hairpin structure and subsequently processed to form open-ended molecules (Roth et al., 1992; Roth et al., 1993). Signal ends and coding ends are subsequently joined to generate heterogeneous coding joints and precise signal joints by slightly different mechanisms (Stanhope-Baker et al., 1996). The coding joint is produced following association of essential proteins with the hairpin structures of Vκ and Jκ segments; the heterodimer Ku70:Ku80, DNA-dependent protein kinase (DNA-PK), the nuclease Artemis, DNA ligase IV and the DNA repair protein XRCC4 (Murphy. K, 2008). The signal joint is the result of association of Ku70:Ku80 and subsequent joining of the signal ends by DNA ligase IV and XRCC4. The imprecise composition of the coding joints

represents an additional layer of diversity embedded within the recombination process. This process is illustrated in Figure 1.5.

Further diversity is introduced to the rearranging Ig genes by the action of terminal deoxynucleotidyl transferase (TdT) which adds non-templated (N) nucleotides to the junctions of the antigen receptor chains. TdT-mediated receptor diversification was initially only associated with the *IGH* genes due to the patterns of TdT expression in the BM, acting only at the pro-B cells stage and down-regulated at the developing pre-B cell stage where light chain rearrangement is initiated, though evidence has been provided for differential TdT activity at both the *IGK* and *IGL* loci in CD45(B220)$^{+}$CD43^{+} pro-B cells from μMT mice (Bentolila et al., 1999).

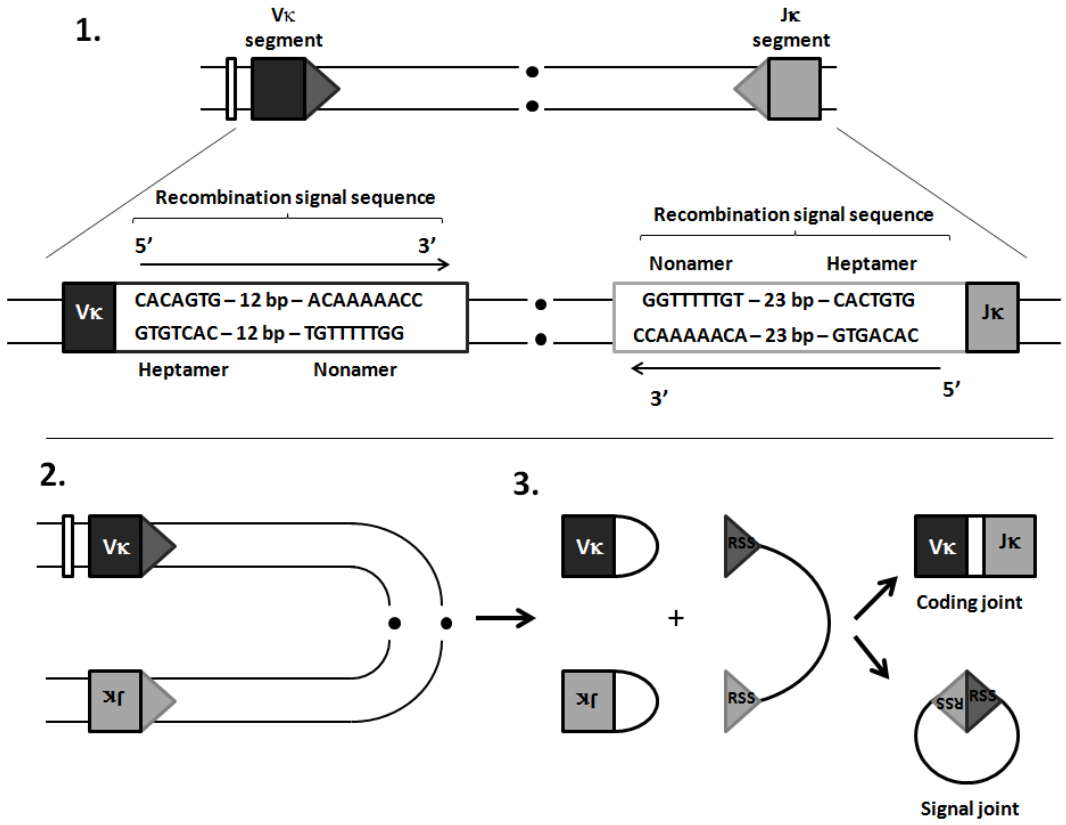


Figure 1.5 V(D)J recombination. **1.** The composition of the recombination signal sequences (RSSs) associated with a Vκ and Jκ gene segment. RSSs contain both a heptamer and nonamer sequence, separated by a spacer region of either 12 or 23 base pairs (bp). Recombination typically takes place between gene segments that are associated with RSSs of opposing orientation and different spacer length. **2.** Recombination is initiated by bringing together the RSSs associated with the recombining gene segments so they are arranged in the same transcriptional orientation in the allele. This results in the looping out of the intervening DNA. **3.** Single strand nicking at the gene segment:RSS junction by RAG1 and 2 proteins results in the generation of a 3'-hydroxyl group that attacks the un-nicked DNA strand forming a hairpin loop. This generates two DNA intermediates. The coding joint is produced following association of Ku70:Ku80, DNA-PK and Artemis, DNA ligase IV and XRCC4 with the hairpin structures of Vκ and Jκ segments. The signal joint is produced following association of Ku70:Ku80 and subsequent joining of the signal ends by DNA ligase IV and XRCC4.

1.1.5 Allelic exclusion

B and T lymphocytes of healthy individuals are genetically committed to express a single antigen receptor with unique specificity. This monoclonal Ig expression is considered to

guarantee a high concentration of a single receptor by these cells. This is paradoxical considering that the hierarchical nature of Ig receptor development can leave a single lymphocyte with the genetic potential to express many simultaneously, and the mammalian immune system is required to respond to vast numbers of foreign antigens with a myriad specificities. The requirement for allelic exclusion and the mechanisms that serve to maintain it will be discussed in the following sections.

1.1.5.1 B cell development and allelic exclusion

Allelic exclusion is achieved by silencing of one of the two homologous lymphocyte receptor alleles. Ig gene rearrangements at each of the three Ig loci in the developing human B cell takes place temporally and sequentially, starting at the activated *IGH* loci and subsequently rearranging genes at the *IGK* then *IGL* light chain loci. Allelic exclusion is exhibited at both the *IGH* and *IGL* loci, and in addition isotypic exclusion is a feature of *IGL* rearrangements, where either a Igκ *or* Igλ light chain will be expressed (Alt et al., 1981; Alt et al., 1984; Ehlich et al., 1993; Korsmeyer et al., 1981). Several mechanisms govern this process, and play a role before and after the recombination events. In order to provide context for the requirement and establishment of allelic and isotypic exclusion, an overview of primary B cell development will be described and the Ig gene rearrangements initiated at each stage of differentiation.

Development of the primary B cell repertoire begins upon differentiation of a non-committed haemopoietic stem cell (HSC) into an early pro-B cell. This marks the beginning of *IGH* gene rearrangement, where D_H-J_H joining takes place simultaneously on activated *IGH* alleles (Ghia et al., 1998). The following V_H- to D_HJ_H recombination event takes place asynchronously on just one allele, where only if the resulting exon produced is non-

productive is rearrangement continued on the second allele. Non-productive rearrangements are those that carry a stop codon, are incompletely rearranged or are in the incorrect genetic reading frame and consequently are not able to produce functional protein (Vettermann and Schlissel, 2010). At this time, the *IGK* light chain loci remain genetically silent and a V_H- to D_HJ_H second rearrangement is attempted on the second allele of the *IGH* locus (Bergman, 1999).

Successful generation of the pre-BCR after pairing of a productively rearranged $V_HD_HJ_H$ exon with surrogate light chains VpreB and λ5 protein represents a fundamental checkpoint for selection in the development of a new B cells (Meffre et al., 2000a), and only successful pre-BCR signalling subsequently permits the initiation of light chain rearrangements as cells that cannot signal through their pre-BCR are induced to die by apoptosis (Keyna et al., 1995; Osmond et al., 1994; Pillai and Baltimore, 1987; ten Boekel et al., 1998). Pre-BCR surface expression also involves interaction with the signalling chains Ig-α/Ig-β (Tsubata and Reth, 1990) and is dependent upon IL-7R signalling. It has been shown in mice that disruption of IL-7R signalling prevented expression of the pre-BCR in new large pre-B cells. Furthermore, successful expression of a pre-BCR lowered the activation threshold of IL-7-induced proliferation, and that there was a positive correlation between IL-7-dependent pre-BCR expression and the activation of the ERK/MAP kinase pathway, which transduces the activation signal downstream of the pre-BCR (Corcoran et al., 1998; Corcoran et al., 1996; Fleming and Paige, 2001). The transition to a small pre-B cell is defined by initiation of primary light chain gene rearrangement at the *IGK* locus (Reth et al., 1987). In the pre-B cell, $V_κ$-$J_κ$ joining is undergone where a productively rearranged exon is then paired with the new and productively rearranged IGHμ protein and expressed on the cell surface as a newly rearranged BCR. This allows exposure to endogenous stromal ligand of the BM as a large pre-B cell. Failure to productively rearrange both alleles at *IGK* loci results in light chain rearrangement proceeding at the *IGL* locus (Hieter et al., 1981).

Cells that fail to productively rearrange an allele at either the *IGH* and both *IGK* and *IGL* light chain loci are removed from the developing B cell pool by apoptosis. Despite the stringent selection checkpoints in the BM, the production of B cells expressing polyreactive and self-reactive receptors is frequent (Wardemann et al., 2003), so the mechanisms used to extinguish these specificities and in many instances rescue the cell from apoptosis will be discussed later. Figure 1.6 outlines the hierarchical heavy and light chain Ig gene rearrangements that take place at different developmental stages of the B cell.

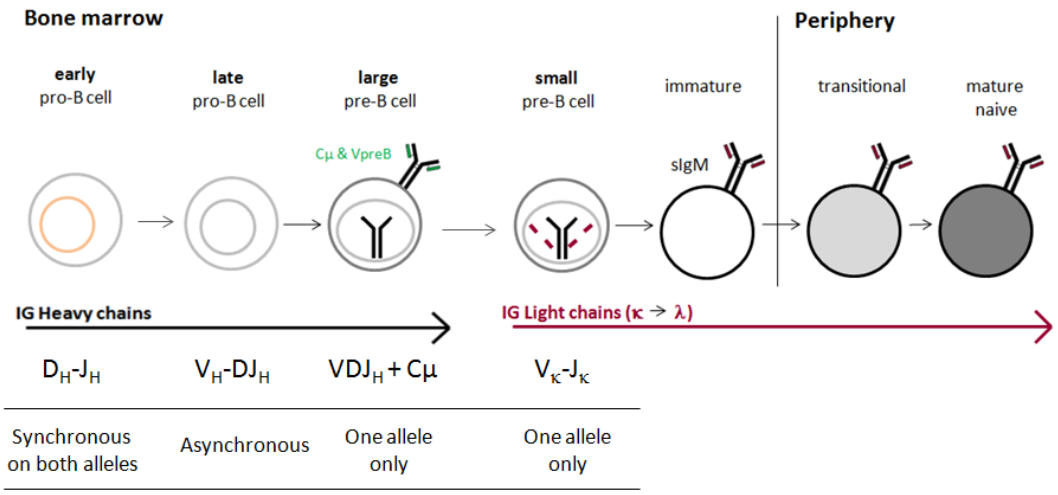


Figure 1.6 B cell gene rearrangements during early B cell development. D_H-J_H rearrangement takes place on both alleles simultaneously in the early pro-B cell. The first *IGH* allele to complete a D_H-J_H rearrangement then undergoes rearrangement with a V_H gene segment as a late pro-B cell. Following complete $V_HD_HJ_H$ recombination the new heavy chain is required to pair with surrogate light chain VpreB and λ5 for expression of a pre-BCR on the cell surface as a large pre-B cell. If this process is successful and a functional heavy chain has been produced, Ig rearrangement at the light chain kappa loci is initiated. At this stage the small pre-B cell undergoes rearrangement sequentially on the first kappa allele (Vκ-Jκ) and the second if rearrangement of the first allele is non-productive. Only if rearrangement at the *IGK* locus is unsuccessful is rearrangement of the *IGL* locus initiated. A successful light chain rearrangement results in the pairing of the new light chain with the heavy chain and expression as a surface IgM BCR. The new B cell can then enter the periphery providing it is not eliminated by mechanism of central tolerance.

1.1.5.2 Models of allelic exclusion and regulation of Ig gene rearrangements

It is acknowledged that the ordered recombination of Ig genes in both humans and mice is not *per se* required for the establishment of allelic exclusion, since other species such as chickens exhibit allelically excluded lymphocytes but rearrange their heavy and light chain genes simultaneously (Ratcliffe, 2006). However, it is clear that allelic exclusion is obeyed during the development of healthy human B cells, and several models can be used to describe its establishment; the asynchronous recombination models, the stochastic model, the feedback inhibition model and the allelic communication model.

The asynchronous recombination models explain how the establishment of allelic exclusion is a result of the restricted chromatin accessibility to Ig alleles and gene segments at a given time during B cell development. Chromatin accessibility and the mechanisms governing nuclear re-localisation of Ig genes and the histone modifications involved in this process are fundamental features of allelic exclusion so will be discussed independently in a later section. There are two variations on the asynchronous recombination model; the probalistic and instructive models. The probabilistic model supports the notion that it is the restricted access to chromatin that serves to limit the frequency of recombination events on one allele. As a consequence of this, the incidence of biallelic recombination events would be improbable (Coleclough et al., 1981; Liang et al., 2004). This model is similar to the stochastic model that proposes that Ig gene rearrangement is maximally efficient, but the probability of productively rearranging more than one Ig allele that encodes a pairing Ig chain is statistically low (Coleclough et al., 1981). These two models differ in that they are each based on a different concept of probability.

The instructive model attributes asynchronous recombination to the earliest replicating Ig allele established during early embryogenesis. This allele is thought to have a higher probability of undergoing rearrangement first, where monoallelic epigenetic changes may take place even before rearrangement is initiated. This hypothesis suggests that the process of allelic exclusion is more similar to X chromosome inactivation than previously considered, where only if the first replicating allele rearranges non-productively is the second allele activated (Mostoslavsky et al., 2001).

The classical model of allelic exclusion is referred to as the feedback inhibition model and postulates that the protein intermediates produced during Ig gene rearrangement serve to restrict and shape the subsequent rearrangement decisions made by the developing B cell (Alt et al., 1980; Alt et al., 1984; Yancopoulos and Alt, 1986). A primary example of this is the contribution of membrane bound IGHμ in the maintenance of allelic exclusion, where membrane bound IGHμ protein (and not soluble IGHμ) has been shown to suppress further heavy chain gene rearrangement (Nussenzweig et al., 1987; Schatz et al., 1992; Storb et al., 1988) and initiates rearrangements at the light chain loci by redirecting the recombinase to the *IGK* locus (Reth et al., 1987; Storb et al., 1988). Targeted disruption of the membrane exon of the IGHμ chain resulted in loss of allelic exclusion, supported by the demonstration that mIGHμ expression is accompanied by decreased transcription through the germline V_H locus (but not the DJ_H locus) and decreased levels of RSS breaks associated with DJ_H alleles (Schlissel and Morrow, 1994).

The allelic communication model is a more recent concept since it is now understood that homologous alleles physically interact and communicate (Vettermann and Schlissel, 2010). It has been demonstrated that Ig alleles pair in a largely RAG-dependent manner, and introduction of DNA breaks to one of the Ig alleles resulted in the activation of ATM,

which acted in trans to reposition the second, uncleaved allele to pericentromeric heterochromatin. It was suggested that this served to prevent biallelic recombination and the potential for chromosomal translocations (Hewitt et al., 2009).

1.1.5.3 Chromatin accessibility and allelic exclusion

Of primary importance to the process of Ig gene rearrangement is the accessibility of V(D)J recombinase to the RSSs within chromatin, which in turn is reliant upon the architecture remodeling, methylation status and histone modifications of the hundreds of gene segments at Ig loci in the nucleus at the specific times during of V(D)J recombination (Schatz and Ji, 2011). The significance of chromatin accessibility to the maintenance of allelic exclusion was demonstrated by Yancopoulos *et al.*, 1985, who observed high expression levels of unrearranged *IGHV* genes specifically in early B cells rearranging V_H-D_HJ_H segments (Yancopoulos and Alt, 1985). Such unrearranged *IGHV* gene expression was not detected in mature Ig-secreting cells. It was subsequently proposed that these findings were the consequence of inactivation of a previously transcriptionally active *IGH* locus – a phenomenon termed germline or sterile transcription. This work was supported by several reports demonstrating the correlation between chromatin accessibility and rates of V(D)J recombination (Schlissel and Baltimore, 1989; Schlissel et al., 1991). It has also been demonstrated that accessibility of the *IGK* locus to V(D)J recombinase and subsequent activation of the *IGK* gene is enhanced by the presence of IGHμ protein following analysis of isolated nuclei *in vitro* (Stanhope-Baker et al., 1996). The same study proposed that the accessibility and cleavage of RSSs is dependent upon the maturity of the cells the nuclei are isolated from, since the V_H locus could be cleaved *in vitro* in the presence of intact pro-B cell nuclei or purified gDNA from pro- or mature B cells, whereas in the nuclei from mature IgD^+ B cells, RSSs flanking V_H gene segments were not cleaved even in the

presence of recombinase. This suggests that chromatin accessibility at the *IGH* locus at a given developmental stage governs heavy chain allelic exclusion (Stanhope-Baker et al., 1996).

1.1.5.3.1 Non-coding RNAs in Ig gene regulation

In the last decade novel studies have exposed a role for non-coding RNA in the field of gene regulation by identifying that a large proportion of the genome is responsible for the production of non-coding RNA transcripts. Alongside understanding of sterile transcription, the genic and intergenic transcription of antisense DNA throughout the *IGHV* locus has been identified (Bolland et al., 2004). Using RNA fluorescent *in situ* hybridisation (RNA-FISH), it was identified that transcription of non-coding antisense DNA strands was strictly regulated and only observed during D_HJ_H to $V_HD_HJ_H$ recombination transition, being down-regulated on both alleles following productive rearrangement on one of the alleles. This led to the hypothesis that this serves as a mechanism to assist in the remodeling of V_H region chromatin at this developmental stage of the B cell (Bolland et al., 2004).

1.1.5.3.2 Histone modifications in Ig gene regulation

Of further importance to the control of allelic exclusion is the covalent modification of histones that mediate chromatin accessibility through methylation and acetylating events. Inactive *IGK*, *IGL* and *IGH* C regions are DNA-methylated, and H3K9 methylation has been proposed to regulate the silencing of Ig loci before the initiation of recombination (Corcoran, 2005; Storb and Arp, 1983). Despite this, it has also been shown that loss of methylation of the *IGK* locus was not sufficient to activate recombination in a recombination-active B cell line with Cre-lox mediated loss of Dnmt1, an enzyme that plays a key role in the establishment of DNA methylation (Cherry et al., 2000).

1.1.6 Isotypic exclusion of Ig light chains

The human immunoglobulin B cell repertoire typically expresses Ig kappa (Igκ) and Ig lambda (Igλ) light chain classes at a ratio of approximately 60% Igκ to 40% Igλ (Hieter et al., 1981; Korsmeyer et al., 1981; Korsmeyer et al., 1982). This differs in mice however, where 95% of B cells express Igκ light chains and just 5% express Igλ (Bräuninger et al., 2001). Studies of light chain rearrangements in cases of human leukemia provided some of the first evidence showing that *IGK* rearrangement proceeds on both alleles before rearrangement of the *IGL* light chains. It was observed that only 2 - 3 % of Igκ-expressing cells harboured a $V_\lambda J_\lambda$ rearrangement joint, whereas nearly all Igλ-expressing cells contained $V_\kappa J_\kappa$ joints (Bräuninger et al., 2001). Exceptions to the ordered rearrangement of Ig genes have been described in both mouse and human B cells, where EBV-transformed pre-B cells isolated from human fetal bone marrow were found to express Igκ light chains without detection of functional IGHμ protein. Of four Igκ-producing EBV-transformed lines established, cells had either just undergone D_H-J_H segment rearrangements, or they carried *IGH* alleles in germline configuration. Despite these observations, exceptions to the ordered light chain rearrangement process are rare and isotypic exclusion is robust in healthy individuals. The observation that single B cells may express more than one Ig light chain, known as allelic inclusion, will be described in a later section.

1.1.6.1 The kappa deleting element

A fundamental mechanism involved in the maintenance of light chain isotypic exclusion involves rearrangements of the kappa deleting element (KDE). The first description of a regulatory element mediating the functional silencing of *IGK* light chain rearrangements in

human B cells followed the observation that just one of twenty rearranged *IGK* alleles isolated from ten Igλ-producing B cells retained its constant region (Hieter et al., 1981).

KDE rearrangements were characterised through mapping and sequencing the 5' rearrangements on both kappa alleles of the Nalm-6 cell line, generated from pre-B cells isolated from a case of acute lymphoblastic leukemia (Graninger et al., 1988). The human KDE, or recombining sequence (RS) in mice, is located 24-25 Kb downstream of the C_κ region (Durdik et al., 1984; Graninger et al., 1988; Siminovitch et al., 1985). Rearrangement of the KDE can occur via one of two pathways; in 75% of instances, the KDE undergoes RSS-mediated site-specific recombination with the palindromic sequence CACAGTG, referred to as the intronic recombination signal sequence (iRSS) located between the distal J_κ and C_κ gene segments (Graninger et al., 1988; Klobeck and Zachau, 1986; Siminovitch et al., 1985). The alternative KDE rearrangement takes place to an upstream, unrearranged V_κ gene segment. Both RSS-mediated recombination events result in deletion of the C_κ region and the enhancers flanking the KDE itself; iE_κ and $3'E_\kappa$, though in the latter instance the V_κ-J_κ joint is also removed (Figure 1.7).

The reason for the different rearrangement sites of the KDE is unknown, though it was initially suggested that RS rearrangement was a pre-requisite for rearrangement at the *IGL* locus (Muller and Reth, 1988; Persiani et al., 1987), where RS rearrangement correlated with Igλ production by B cells (Vela et al., 2008). Subsequent studies however demonstrated that Igλ-expressing B cells still are produced in the absence of KDE rearrangement (Zou et al., 1993). Of further interest, when RS rearrangement was prohibited in mouse models, the frequency of sequential light chain gene rearrangements (receptor editing) were impaired in response to autoantigen exposure and the frequency of Igλ-producing B cells observed was reduced (Vela et al., 2008).

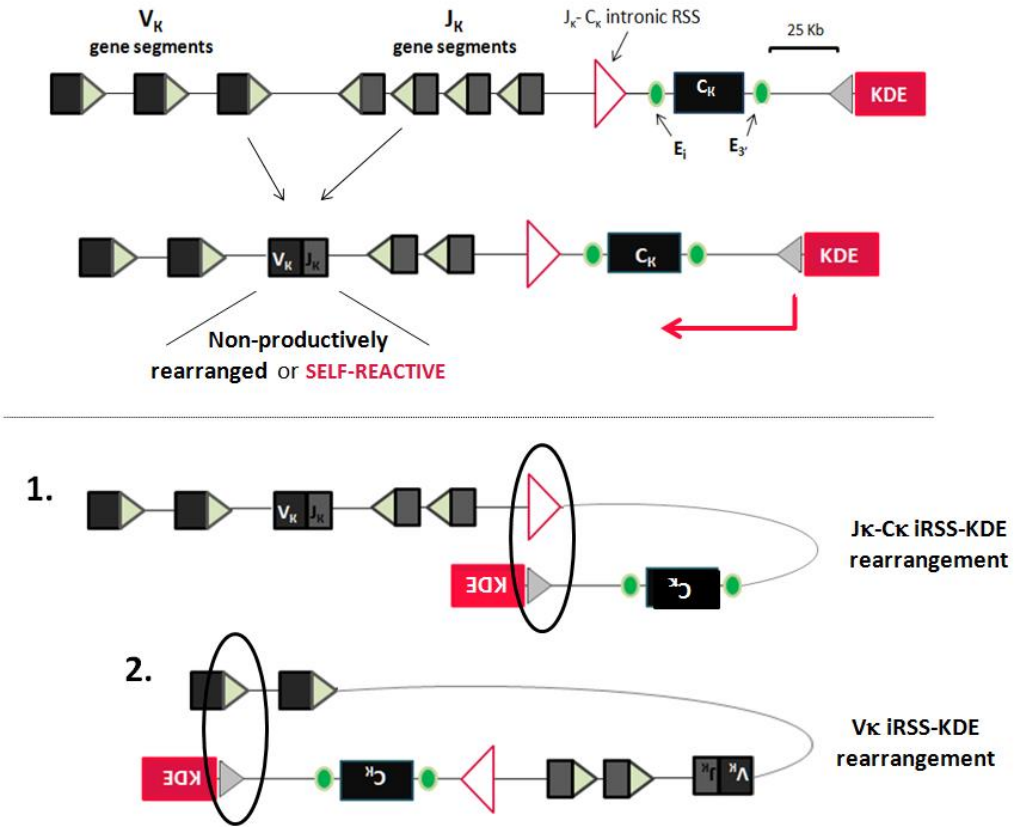


Figure 1.7 KDE rearrangements. The KDE is located 25 kb downstream of the Cκ gene segments and is flanked by two enhancers; iEκ and 3'Eκ. Recombination of the KDE takes place when a Vκ-Jκ rearrangement is either non-productive, or generates a self-recognising specificity. The KDE rearranges by two different mechanisms – it involves either **1**. rearrangement to an intronic RSS located between the Jκ and Cκ gene segments deleting Cκ and the enhancers flanking it (green ovals), or **2**. to an RSS of an upstream and unrearranged Vκ segment deleting the rearranged VκJκ joint in addition to the 3' Cκ region.

The functional consequences of KDE (RS) rearrangements and the role of the enhancers iEκ and 3'Eκ and the constant region C_κ in mediating *IGK* locus silencing were elucidated using transgenic mice carrying neomycin resistance genes in the place of C_κ or the iEκ. Studies have described that mice homozygous for both C_κ and iEκ inactivation produce 7-fold more Igλ-expressing B cells (and around 2.3 times fewer total B cells) compared to the WT mice. Heterozygous mutants produce different phenotypes; iEκ inactivation completely blocks rearrangement at the *IGK* locus whereas Cκ inactivation alone permits Vκ-Jκ rearrangement though rearrangements of the RS in these mice are

lower in frequency than in WT mice (Zou et al., 1993). Additional studies have shown that the 3'E□□enhancer is required for high levels of *IGK* gene expression, and it has been suggested that this enhancer alone is sufficient for the transcription of *IGK* genes in mature B cells (Meyer and Ireland, 1998; Takeda et al., 1993). Interestingly, double mutant B cells for the iEκ and 3'Eκ enhancers were unable to rearrange their *IGK* genes, but could rearrange and subsequently express *IGL* light chains (Inlay et al., 2002).

The enhancers flanking the KDE have also been implicated in conferring the ability to acquire mutations within *IGK* alleles. Analysis of Ig gene rearrangements and somatic mutation in monocytoid B cell lymphoma (MBCL) Igλ-expressing cells demonstrated that V_H genes were highly mutated, whereas *IGK* genes that had rearranged and were KDE inactivated were unmutated (Kuppers et al., 1996). Of note, both productive and non-productively rearranged heavy and light chains can undergo mutation (Kocks and Rajewsky, 1989). These data suggested that the deleterious KDE recombination event prevented B cells from acquiring somatic mutations in their *IGK* rearranged genes if the cell were to become activated peripherally (Betz et al., 1994). These data are important since not only non-productively rearranged *IGK* alleles are KDE inactivated. An elegant study that involved Ig sequence analysis of gDNA isolated from FACS-sorted single naive Igκ and Igλ-producing B cells identified that 30% of all Vκ-Jκ rearrangements isolated from Igλ-expressing B cells were actually productively rearranged. These findings suggested that despite being in the correct genetic reading frame, a molecular decision was made to not use these productive rearrangements of *IGK* and inactivate them by rearrangement of the KDE (Bräuninger et al., 2001).

The frequency of KDE rearrangements in populations of B cells has been suggested to represent an indicator of light chain rearrangement activity (Panigrahi et al., 2008). KDE

rearrangements also take place independently of antigen specifity and do not encode functional protein, reinforcing the potential for using KDE rearrangement status as a marker of extensive light chain rearrangements (Vela et al., 2008).

A prominent study by Panigrahi *et al.*, 2008 studied the rearrangement frequency of the KDE to the iRSS in health and SLE by isolating the gDNA from FACS-sorted populations of both Igκ- and Igλ-producing $CD19^+$ B cells and performing real-time PCR anaysis for quantification of the iRSS-KDE rearranged junction (Panigrahi et al., 2008). They observed that KDE rearrangements were significantly decreased in frequency in SLE compared to controls. Reduced KDE rearrangement frequency was also observed in Igλ-producing $CD19^+$ B cells from type I diabetic patients (T1D), though this was not observed in Igκ-producers. In light of this data it was proposed that the silencing of *IGK* light chain rearrangements may be defective in SLE and T1D, interpreted as a defective ability to edit light chains; a mechanism of B cell tolerance. Studies of the murine equivalent of the KDE, the RS, have also provided evidence for a role for the KDE in secondary gene rearrangements, since RS knockout mice produce Igλ-expressing B cells with self-specificity in the periphery (Vela et al., 2008).

1.1.7 B cell tolerance

1.1.7.1 Mechanisms of B cell tolerance

The importance of establishing self-tolerance is critical to mammalian immune homeostasis to avoid what is referred to as 'horror autotoxicus', a termed coined by Paul Ehrlich at the turn of the century and represented the earliest understandings of autoimmunity (Steinman and Nussenzweig, 2002). Studies using transgenic mice have elucidated roles for three main mechanisms that mediate self-tolerance and the removal of

autoreactive specificities in the development of the primary B cell repertoire in the BM; clonal deletion, the induction of anergy and receptor editing (Casellas et al., 2001; Radic et al., 1993).

Clonal deletion was first described following isolation of BM cells from transgenic mice for Ig rearrangements encoding anti-H-2K^k antibodies and subsequent injection into lethally irradiated mice expressing or not expressing the H-2K^k antigen. It was shown that transgenic mice expressing the H-2K^k antigen had deleted H-2K^k-specific B cells from the spleen and LNs, whereas mice lacking H-2K^k antigen carried an abundance of H-2K^k-specific B cells in circulation (Nemazee and Buerki, 1989). This mechanism was mediated by receptor cross-linking in immature B cells (Nemazee and Weigert, 2000). Strong signals via the BCR are thought to induce immature B cell deletion by apoptosis (Meffre and Wardemann, 2008).

Some prominent studies in the late 1980s used transgenic mice exposed to self-antigen and observed down-regulation of surface IgM in response to BCR stimulation that exceeded the threshold for activation. This appeared to correlate with clonal silencing of self-responsive B cells and they proposed that a proportion of the circulating peripheral B cell repertoire were tolerised by this mechanism (Goodnow et al., 1988; Goodnow et al., 1989; Nossal and Pike, 1980). This is known as the induction of anergy. A recent study has described the identification of a population of human B cells that are predominantly autoreactive, as demonstrated by binding to HEp-2 antigens, but tolerised by functional attenuation. These cells were suggested to be fully mature IgD$^+$IgM$^-$ B cells, referred to as B_{ND} cells, and are thought to represent 2.5% of peripheral blood (Duty et al., 2009).

Receptor editing can rescue a newly rearranged light chain that is self-reactive for ubiquitous membrane-bound autoantigens of the BM stroma by a sequential and nested rearrangement event on the same allele, altering the specificity in order to eliminate self-

reactivity (Halverson et al., 2004). The anti-DNA antibodies spontaneously produced in MRL/lpr mice aided in discovery of receptor editing as a mechanism to achieve tolerance. B cells of adult transgenic mice bearing both the *IGH* and *IGL* chain of anti-DNA antibodies only expressed the transgenic *IGH* paired with endogenous light chains using a restricted repertoire of Vκ segments. That was the case despite detection of transcription of the transgenic, dsDNA-specific light chain. These observations suggested that autoreactive transgenic B cells may alter their light chains to enhance the likelihood of producing a non-self reactive receptor and consequently escape deletion (Gay et al., 1993). Receptor editing takes place at the light chain loci following intermediate levels of BCR antigen engagement in the developing B cell (Meffre and Wardemann, 2008). Ig heavy chains do not have the capacity to undergo receptor editing, since the process of successful V-D_HJ_H rearrangement removes remaining D_H gene segments from the repertoire, and recombination cannot take place between distal V_H and J_H segments due to the same orientations of their RSSs. It was proposed that receptor editing is the primary mechanism used to rescue autoreactive B cells, where clonal deletion only becomes necessary when no unrearranged gene segments are available on the same light chain alleles to permit further editing (Halverson et al., 2004). Furthermore, a significant shift in J_κ gene segment usage between the $V_\kappa J_\kappa$ joints of Igκ and Igλ-expressing B cells has been observed, indicating that Igλ-producers undergo sequential $V_\kappa J_\kappa$ rearrangements on the same allele before undergoing rearrangement at the lambda allele (Bräuninger et al., 2001).

Sequential rearrangements events of the light chain loci in peripheral lymphoid organs is often referred to as receptor revision and is a more contentious concept, though studies have demonstrated that secondary rearrangements of Ig light chains are commonly undergone by IgA^+ plasma cells in a healthy gut (Su et al., 2008). The identification of biases in the productive and non-productive repertoires of Igλ-expressing B cells, evidence of *RAG* gene and protein expression and identification of rearrangement intermediates

support this hypothesis (Su et al., 2004). In addition, T-cell dependent replacement of light chains has been described to take place in the germinal centres (GCs) of secondary lymphoid tissues to further diversify the Ig repertoire and select for high affinity BCRs and eliminate autoreactive receptors generated by somatic hypermutation.

Studies in mouse models have similarly identified products of recombination and the protein product of RAG1 and -2 genes in GCs (Hikida et al., 1998). The expression of RAG genes is a much-debated subject, with evidence for and against its re-expression in peripheral lymphoid tissues. Although *RAG1* and *RAG2* gene expression is typically thought to occur in two waves during B cell development, catalysing *IGH* chain rearrangements followed by recombination at the *IGL* loci, it has been reported that RAG1 and -2 are re-expressed in human GC B cells that have lost surface Ig, suggesting the ability to undergo a secondary round of Ig gene rearrangement (Giachino et al., 1998). In addition, the detection of RAG has been reported in activated B cells isolated from human tonsil (Girschick et al., 2001). Of note, the reactivation of *RAG* gene expression in mature lymphocytes has been suggested to be responsible for aberrant rearrangement events involving Ig or TCR loci and proto-oncogene loci (Korsmeyer, 1992). Evidence against the re-induction of *RAG* used mice carrying GFP-tagged *RAG* gene products and illustrated in their hands that *RAG* gene expression is retained within the transitional B cell compartment of peripheral blood, proposing that re-induction of these genes in peripheral tissues does not take place (Nagaoka et al., 2000).

1.1.7.2 B cell tolerance checkpoints

As described above, clonal deletion, the induction of anergy and receptor editing exist to maintain central tolerance in the BM by negative selection of autoreactive specificities. The actual frequencies of self-reactive antibodies generated during the development of B cells

had previously not been studied until 2003 when Wardemann *et al.*, 2003 investigated the development and silencing of self-reactive B cells in three healthy subjects. Individual antibodies from B cells at different stages of development were isolated from peripheral blood and BM and their antibodies were expressed in 293A cell lines. A total of 248 antibody clones were then tested for binding to the HEp-2 cell line extract – a standard clinical assay for the detection of true ANAs that may be pathogenic and are hallmark of SLE (Leslie et al., 2001; Wardemann et al., 2003).

Enzyme-linked immunosorbent assays (ELISAs) demonstrated that 75.9% antibodies cloned from early immature B cells showed high levels of reactivity in this assay, which decreased to 43.1% in B cells at the immature stage. HEp-2 reactivity remained at 40.7% in newly emigrated B cells in the blood, but an additional drop in reactivity in was observed between the new emigrant and naïve B cell compartments (Wardemann et al., 2003). Of the 75.9% antibody clones from early immature B cells, 58.6% showed nuclear or nuclear/cytoplasmic staining by indirect immunofluorescence assays (IFAs) on fixed HEp-2 cells. Long *IGH* CDR3 regions have previously been associated with self-reactive antibodies (Aguilera et al., 2001), and this study illustrated that cloned antibodies from early progenitor B cells had significantly longer CDR3 regions than the mean CDR3 length of *IGH* antibody chains from peripheral B cells. Positively charged CDR3s were also enriched within pre-B and early immature B cell pools. Both these features were selectively removed as B cells developed. These data demonstrate in healthy individuals, self-reactive antibodies are eliminated from the B cell repertoire at the early immature to immature B cell stage, representing a checkpoint for B cell tolerance (Wardemann et al., 2003).

1.1.7.3 Allelic inclusion – failure of B cell tolerance?

The ability to express more than one rearranged Ig allele has been demonstrated in genetically modified mice engineered to develop a B cell repertoire which stably expressed two different heavy chains (Sonoda et al., 1997). It has also been suggested that the co-expression of a non-self-reactive Ig light chain can rescue autoreactive B cells from negative selection by 'diluting' the self-reactive receptor, and essentially 'bypassing' central tolerance (Gerdes and Wabl, 2004; Li et al., 2002b; Liu et al., 2005). An examination of how tolerance to autoantigen is established involved the K/BxN mouse model of spontaneous and inflammatory autoimmune arthritis. Arthritogenic B cells expressing antigen-specific B cell receptors (BCRs) with different affinities for the autoantigen glucose-6-phosphate isomerase (GPI) were used to generate knock-in mice for in which the rearranged *IGH* and *IGL* variable gene segments of an anti-GPI hybridoma were targeted into the corresponding germ-line J loci. It was observed that sequential light chain rearrangements on the same allele, receptor editing, contributed to the removal of autoreactive BCRs with a high affinity for GPI. It was also suggested that the expression of an additional light chain by some of the GPI-reactive B cells contributed to their 'escape' from negative selection due to their low affinity for antigen (Huang et al., 2006). Interestingly, these findings have led to the proposal that light chain allelic inclusion may represent an additional B cell tolerance mechanism (Makdasi and Eilat, 2013).

It is thought that receptor editing compromises allelic exclusion since it sustains ongoing recombination events even after production of a functional Ig receptor (Li et al., 2002b). One study used gene targeting to replace the mouse *IGKC* region with the human counterpart allowing for the generation of normal mice heterozygous for the human *IGKC* allele ($Ig\kappa^{m/h}$) in the context of a wild-type repertoire. Ten percent of mature B cells were found to co-express both alleles of *IGK* and that these cells participated in immune

responses (Casellas et al., 2007). Thirty percent of these dual-light chain expressing B cells exhibited autoreactivity in screens for dsDNA or Hep-2 cell binding. They also suggested that light chain inclusion by these cells was the result of receptor editing since usage of *IGKJ5* was significantly enriched within the dual-expressing B cell population (Casellas et al., 2007). Furthermore, Igκ^{h} mice bearing a gene targeted human IGKC allele were crossed with the lupus-prone MRL-*Fas*$^{lpr/lpr}$ in a study that demonstrated that B cells expressing two Igκ light chains were increased in frequency with the age and disease progression of the mice. In addition, the dual-Igκ -expressing B cells exhibited higher levels of autoreactivity than single Igκ cells and were frequently observed to have undergone antigen activation and differentiated into memory and autoantibody-secreting plasmablast populations (Fournier et al., 2012).

A very recent study however has provided evidence to contradict this study, finding that just 0.5% of anti-DNA-producing hybridomas from knock-in lupus-prone NZB/NZW mice appeared to exhibit light chain inclusion, and that those that do express allelically included light chains do not express higher proportions of autoreactive BCRs. Ultimately, whether the allelic inclusion of light chains is a protective tolerance mechanism, or the by-product of ongoing light chain gene rearrangements, it remains to be a highly interesting area of study.

1.1.8 Human B cell subsets

The development of human B cell subsets can largely be defined by the status of their immunoglobulin genes, and those that have successfully undergone gene rearrangement events on both the *IGH* and *IGL* loci in the BM enter the peripheral circulation. The peripheral development of B cells can be broadly separated into five sequential stages;

transitional B cells that represent new BM emigrants, mature naïve B cells that are antigen inexperienced, germinal centre B cells that are engaged in immune responses to cognate antigen, memory B cells that survive for prolonged periods and plasma cells that secrete soluble antibodies (Bemark et al., 2012).

1.1.8.1 Differences between mouse and human B cell development

Human B cell development is not equivalent to the development of the murine B cell repertoire (Weill et al., 2009). During murine B cell development, immature transitional B cells are thought to be comprised of four transitional B cell subclasses; T0, T1, T2 and T3, defined by their expression of CD23 and IgM and AA4, a fetal stem cell marker (Allman et al., 2001). T2 B cells are thought to be derived from T1 B cells and are located predominantly in the spleen, and T3 B cells are considered to represent an anergic B cell population (Teague et al., 2007). Upon reaching the spleen murine transitional B cells become mature naive B cells existing as either splenic resident marginal zone (MZ) or circulating follicular (FO) B cells; these subsets thought to be specialised for taking part in T-independent and T-dependent immune responses respectively (Garraud et al., 2012; Vossenkamper et al., 2012). Murine B1 cells represent a fetally derived B cell lineage and are restricted to the peritoneal and pleural cavities and are a self-renewing population (Carey et al., 2008). There is no equivalent B cell subset in humans and the anatomical differences between the mouse and human MZ supports the notion that human B cell development is unlikely to occur by the same pathway (Mebius and Kraal, 2005).

1.1.8.2 Human transitional B cells

Human transitional B cells comprise approximately 2-3% of human peripheral blood (Vossenkamper et al., 2012) and can be phenotypically identified by the high cell surface

expression of CD19, CD24, CD38 and CD10 (Palanichamy et al., 2009; Sims et al., 2005). Human transitional B cell subsets can be further divided into T1, T2 and T3 populations. The T1 population can be sub-divided into $CD21^{lo}$ or $CD21^{+}$ cells, with T1 $CD21^{lo}$ cells considered to be at a more immature developmental stage than the $CD21^{+}$ T1 population. It has been suggested that T1 $CD21^{lo}$ cells may represent an anergic transitional B cell population since the down-regulation of CD21 expression following antigen encounter has been associated with the induction of anergy (Andrews and Wilson, 2010; Suryani et al., 2010).

Human transitional B cells may also be identified in populations of $CD27^{-}IgM^{+}$ B cells through the lack of expression of the ATP-binding cassette B1 (ABCB1) transporter, since ABCB1 expression is restricted to resting naive B cells (Wirths and Lanzavecchia, 2005). These two methods for identifying transitional B cells led to the designation of the T3 transitional B population that is phenotypically naive, though lacks expression of ABCB1 (Bemark et al., 2012; Palanichamy et al., 2009). The differentiation of a human transitional B cell and its progression to becoming an MZ B cell is poorly understood, but the recent suggestion that transitional B cells enter gut-associated lymphoid tissue (GALT) as part of normal human B cell development has contributed to this area of debate (Vossenkamper, 2013). In this study, T2 cells were demonstrated to be selectively recruited to Peyer's patches of GALT, supported by the highest expression of the α4β7 integrin by this transitional B cell subset. T2 cells were shown to be activated *in vivo* by detection of phosphorylated of BTK, Erk and Syk and approximately 30% of GALT T2 cells had acquired somatic mutations in their *IGVH* genes, indicative of participation in an immune response (Vossenkamper, 2013). These findings suggest a novel, previously un-described pathway of human B cell development and repertoire development in humans. It is

unknown what the fate of these T2 cells may be; they may die in the gut or may be the precursors of either human MZ cells, enigmatic in origin, or intestinal plasma cells.

1.1.8.3 Human naive and memory B cell subsets

The human naive B cell population used to be phenotypically characterised by the expression of unmutated IgD or IgM and absence of CD27 expression, since these B cells represented a population that had not encountered antigen or undergone subsequent CSR to produce either IgG, IgA or IgE (Wu et al., 2011). Naive CD27- B cells comprise approximately 60-70% of the total B cell pool (Weill et al., 2009). However, it is now understood that a significant population of IgM^+IgD^+ cells carry V region mutations and comprise around 15-20% of the total lymphocytes in the peripheral blood of healthy subjects (Dunn-Walters et al., 1995; Klein et al., 1997; Klein et al., 1998; Weill et al., 2009). Since this IgM^+IgD^+ subset also expressed CD27, it was proposed that CD27 was an appropriate marker for the identification of memory B cell subsets (Klein et al., 1998). The origin of this B cell subset remained elusive for some time, though more recent studies have demonstrated that $IgM^+IgD^+CD27^+$ and $IgM^+IgD^{-/lo}CD27^+$ cells both carry mutations in Bcl6, indicative of cells participating in germinal centre responses. In addition, the $IgM^+IgD^+CD27^+$ population were found to be clonally related to GC-derived IgG^+ cells that shared V gene mutations (Seifert and Kuppers, 2009). The identification that naive cells could be discriminated from memory B cell populations by their expression of ABCB1 also aided these studies (Wirths and Lanzavecchia, 2005).

Memory B cells respond more rapidly to antigen encounter than naive B cells with the same specificity and use a different repertoire of *IGVH* gene segments than naive B cells (Blanchard-Rohner et al., 2009; Tangye et al., 2003; Wu et al., 2010). The identification of memory B cell subsets is complex. The early suggestion that CD27 represents a general

marker of all memory B cell subsets (Klein et al., 1998) is being questioned following the recent identification that $CD27^-$ B cells actually represent a heterogeneous B cell pool. A population of IgG^+CD27^- B cells carried somatically mutated V_H genes that indicated germinal centre involvement (Fecteau et al., 2006). A recent study has described that the detection of the CD45RB epitope using a mAb known as MEM55 ($CD45RB^{MEM55}$) is a useful method for identification of all memory B cell subsets independently of CD27 expression (Koethe et al., 2011). Splicing of the CD45RB exon was found to be unchanged during B cell differentiation, and that its expression is regulated by modifications by O-linked glycosylation. CD45 is a highly expressed tyrosine phosphatase that is involved in the regulation of lymphocyte signalling activation strength during antigen encounter (Koethe et al., 2011).

1.1.8.4 Human plasmablasts and plasma cells

The plasmablast B cell pool consists of B cells immediately preceding differentiation into plasma cells (Bemark et al., 2012). Plasmablasts can be phenotypically defined as $CD19^{lo}CD20^{-}CD27^{high}CD38^{high}CD138^{+/-}$ B cells and are predominantly of mucosal origin; 84% have been described to secrete IgA (Caraux et al., 2010; Mei et al., 2009). Expression of CD138, also referred to as syndecan-1, is typically acquired as a marker of both human and mouse plasma cells and is thought to mediate adhesion of plasma cells to the BM stromal matrix via interactions with type I collagen (Jego et al., 2001; MacLennan et al., 2003; O'Connell et al., 2004). Plasmablasts that express similar levels of CD27 and CD38 however are commonly either positive or negative for CD138 expression. This makes distinction between circulating plasmablasts and plasma cells limited (Fink, 2012) and as a result, these cells are commonly referred to as antibody-secreting cells (ASCs) (Wrammert et al., 2008). Very low numbers of plasma cells (2/μL) are found in peripheral blood of healthy subjects (Caraux et al., 2010) though increases in plasmablast and plasma

cell number are observed in circulation following vaccination or infection (Wrammert et al., 2008).

1.1.8.5 Human regulatory B cells

B cells are typically considered to be positive regulators of immunity, largely due to their production of antibodies (Yang et al., 2013). However, the role of B cells with regulatory capacity have been of great interest since identification of the broad range of immunosuppressive cytokines B cells can secrete, such as IL-10 and TGFβ. One of the earliest descriptions of a suppressive role of B cells was proposed in 1974, where studies investigating delayed type hypersensitivity (DTH) skin responses in guinea pigs demonstrated that DTH was impaired just prior to the detection of antibody (Katz et al., 1974; Neta and Salvin, 1974). A regulatory role for B cells in restricting autoimmunity was elucidated from studying mouse models of experimental autoimmune encephalomyelitis (EAE). Increased disease severity and failure to recover from disease was observed in EAE mice deficient in B cell function due to disruption of their IGHμ chain transmembrane exon (Wolf et al., 1996). Similar studies using the EAE mouse model illustrated that B cells producing IL-10 following autoantigen recognition were responsible for the recovery from disease and resolution of inflammation (Fillatreau et al., 2002). This IL-10-producing subset of B cells was also shown to up-regulate cell surface expression of CD1d and produced IL-10 which directly played a role in the suppression of IL-1 and STAT3-induced pathway of inflammation (Mizoguchi et al., 2002). The following year, a role for B cell derived IL-10 in the protection against arthritis development was described (Mauri et al., 2003). This study used the chronic collagen-induced arthritis (CCIA) model, which represents a Th1 driven arthritic disease and develops following immunization of DBA/1-TcR-β-Tg mice with type II bovine collagen (CII) in complete Freund's adjuvant (CFA)

(Mauri et al., 1997). They demonstrated that the activation of arthritogenic splenocytes *in vitro* with agonistic anti-CD40 mAb and CII as the source of antigen induced the development of B cell population producing high levels of IL-10 and low levels of IFN-γ. Adoptive transfer of these cells into DBA/1-TcR-β-Tg mice following immunization with CII inhibited Th1 cell differentiation, prevented the development of arthritis in these mice and could ameliorate established disease (Mauri et al., 2003).

It was later suggested that the origin of these regulatory IL-10 producing B cells could be attributed to a B cell subset expressing high levels of CD21, CD23 and IgM; transitional 2 marginal zone precursor (T2-MZP) B cells (Evans et al., 2007). Adoptive transfer of this B cell subset into immunized DBA/1 mice during the induction phase of CIA ameliorated disease by inhibition of Th1 immune responses via and IL-10 mediated mechanism (Evans et al., 2007). Human B cells exhibiting a regulatory capacity have been described as a population of cells with the surface phenotype $CD19^{+}CD24^{hi}CD38^{hi}IgD^{+}CD1d^{hi}$ that constitute approximately 1% of peripheral blood. Healthy $CD19^{+}CD24^{hi}CD38^{hi}$ cells have been shown to inhibit the differentiation of naïve $CD4^{+}$ T cells into Th1 and Th17 cells partially through IL-10 secretion. In addition, $CD19^{+}CD24^{hi}CD38^{hi}$ cells of healthy donors converted CD4+CD25- T cells into T_{REGS} (Blair et al., 2010; Flores-Borja et al., 2013). $CD19^{+}CD24^{hi}CD38^{hi}$ cell numbers have also been correlated with IL-10 and a state of tolerance in renal transplant patients (Newell et al., 2010). The studies involving investigating regulatory B cells in SLE will be described in a later section.

Several modes of activation of regulatory B cells have been described. Regulatory B cell function is largely dependent upon interaction with effector T cells, where a body of evidence supports the requirement for CD40-CD154 interaction in their activation. EAE mice harbouring B cells with a CD40-specific deficiency suffer from more severe disease and reduced levels of IL-10 production (Fillatreau et al., 2002). The use of monoclonal

antibodies targeting CD40 has demonstrated therapeutic benefit to human subjects suffering from arthritis which appears to redirect the Th1 response by provision of IL-10 (Mauri et al., 2000). TLR stimulation has been described to be the strongest stimulus for B cell-mediated T cell suppression through IL-10 production in both mice and humans. TLR2 and TLR4 are required for regulatory B cell function in mice following LPS activation and relieve EAE (Lampropoulou et al., 2008). In humans, optimal stimuli of B cells to induce IL-10 production involves activation of TLR9 using CpG-B in combination with BCR stimulation using anti-Ig. B cells stimulated in this way partially inhibit the proliferation of $CD4^{+}CD25^{-}$ naïve T cell populations *in vitro* (Bouaziz et al., 2010).

1.1.9 B cell activation and immune homeostasis

Maintenance of B cell homeostasis and the appropriate response following activation requires the integration and regulation of a number of intracellular signals delivered from multiple surface receptors; predominantly the BCR, Toll-like receptors (TLRs), the B cell activating factor (BAFFR) receptor, and CD40, as well as intracellular cues. Failure to integrate pathways such as NF-κB signalling can lead to B cell deficiency, aberrant B cell activity, or even lymphoma (Tavares et al., 2010). Aberrant B cell tolerance and selection can cause production of autoantibodies, formation of immune complexes (ICs) and tissue damage and the manifestation of autoimmune disease states (Fairhurst et al., 2006).

The molecular decisions made following activation of a B cell after antigen encounter determines the fate of the B cell and ultimately the generation of a successful adaptive immune system. The most basic mode for B cell activation and subsequent survival involves the integration of two signals; the binding of cognate antigen to the BCR and the subsequent engagement of co-stimulatory molecules in a T-dependent interaction of TNF-family receptor super-family member CD40 expressed on the B cells with co-stimulatory

molecule CD40L (CD154) on the surface of the cognate T cell. CD40 a member of the TNFR family and is also expressed by professional antigen presenting cells (APCs) and non-immune cells (Elgueta et al., 2009). The strength, duration and B cell developmental stage of CD40 ligation plays a role in shaping whether a mature B cell secretes Ig or whether Ig secretion is suppressed (Miyashita et al., 1997). The second signal may also be provided via TLR stimulation. B cells may be activated by T-cell dependent and independent pathways.

1.1.9.1 T-cell independent B-cell activation

The formation of a germinal centre is dependent upon the function of T_H cells, though the responses of some B cells occur independently of this cognate interaction. T-independent B cell activation can occur through interactions with microbial components, typically large lipid or carbohydrate molecules, and induce strong antibody responses in healthy individuals. These are known broadly as thymus-independent (TI) antigens which can be further classified into TI-1 and TI-2 subgroups depending on the activation mechanism triggered by the stimulating antigen, though both stimulate rapid IgM production in the absence of cognate T cell help (El Shikh et al., 2009).

TI-1 antigens are capable of polyclonally activating mature and immature B cell populations non-specifically resulting in their proliferation and differentiation, providing they are at a high concentration (Murphy. K, 2008). In mice, lipopolysaccharide (LPS) is a TI-1 antigen and a potent B cell mitogen expressed on the outer membrane of gram-negative bacteria where recognition by TLR4 following binding to LPS-binding protein and CD14 elicits an immune response. TI-2 antigens tend to consist of multiple repeating subunits of bacterial capsular polysaccharides; the most potent TI-2 antigens simultaneously cross-link a number of BCRs, and it has been demonstrated that TI-2

molecules must consist of a minimum of twenty repeating epitopes that were appropriately spaced in order to elicit immunogenicity (Dintzis et al., 1976; Dintzis et al., 1983). TI-2 antigens contain no intrinsic B cell stimulating activity and can only activate mature B cells (Fagarasan and Honjo, 2000; Lesinski and Westerink, 2001). B cell activation by TI-2 antigens is supported by BAFF, a co-stimulatory molecule provided by DCs, and these antigens have also been shown to play a role in triggering activation of the complement pathway (Mond et al., 1995). The density of TI-2 antigen is considered to be pivotal to the B cell response induced, where excessive cross-linking of B cell receptors by TI-2 antigens is considered to render the B cell anergic, and low TI-2 antigen concentrations fail to induce its activation (Murphy. K, 2008).

1.1.9.2 T-cell dependent B-cell activation

Following human B cell development in BM, B cells enter the periphery as transitional B cells expressing a distinct cell surface molecule phenotype $CD19^{+}IgD^{+}CD27^{-}CD24^{hi}CD38^{hi}CD10^{hi}$. These new B cells are thought to migrate via high endothelial venules (HEVs) in T cell zones harboring patrolling and unstimulated $CXCR5^{lo}$ and $CCR7^{hi}$ to secondary lymphoid organs such as the spleen, Peyer's patches, lymph nodes (LNs) and tonsils where B cells then migrate to B cell follicles if antigen is encountered (Goodnow et al., 2010). Transit to the follicle is facilitated by the expression of CXCR5 and CCR7 which mediates chemotaxis towards CXCL13-expressing follicular dendritic cells and stromal cells producing CCL19 and CCL21 respectively (FDCs).

B cells that fail to encounter cognate antigen re-enter the peripheral circulation via efferent lymph vessels (Okada et al., 2005; Reif et al., 2002). BCR engagement induces higher expression of CCR7 within 1-6 hours of antigen receptor binding to facilitate migration to the follicle/T cell zone border and expression of CD86 to mediate interaction with CD28

on the cognate T cell. At this location, self-reactive B cells undergo cell death as prolonged antigen exposure down-regulates CD86 expression whereas non-autoreactive B cells begin to proliferate within 1-2 days at this site and reposition themselves to perifollicular areas rich in $CD11c^{hi}$ DCs that are responsible for production of BAFF and APRIL survival signals to the B cell (Steinman et al., 1997). By day 3 molecular decisions determine whether a single proliferating B cell becomes a re-circulating early memory B cell, an extrafollicular plasma cell or a germinal centre (GC) B cell. B cell clones with the highest affinity tend to be selected for extrafollicular plasmablast compartment (Chan et al., 2009) and differential expression of chemoattractant receptors shape the fate of the B cell; for example, plasma cells lose CXCR5 and express high levels of CXCR4 whereas GC B cells retain CXCR5 expression.

1.1.9.2.1 Somatic hypermutation and class-switch recombination

Upon reaching a GC, mature B cells that have encountered cognate antigen migrate towards the dark zone of the GC as rapidly proliferating centroblasts and undergo somatic hypermutation (SHM) and class switch recombination (CSR) upon entering the light zone; both processes occur during antigen-driven differentiation of B lymphocytes and are absolutely dependent upon the expression of activation induced cytidine deaminase (AID) specifically by activated B lymphocytes (Dunnick et al., 2009; Neuberger et al., 2003). AID was identified to be responsible for governing both processes through studies of AID-deficient mice and humans. Mouse B cells specifically express AID in GCs, and AID-deficient mice develop hyper-IgM syndrome and are unable to make IgG, IgA or IgE antibodies. Human studies have revealed that patients with the autosomal recessive form of hyper-IgM syndrome (HIGM2) carry mutations in the human counterpart AID which manifests as an inability to class switch Ig, the inability to acquire somatic hypermutations

and lymph node hyperplasia, thought to be caused by the presence of very large germinal centres (Revy et al., 2000).

AID deaminates deoxycytidine residues to uracil residues in single stranded DNA. The resulting U:G mispairs act as lesions for repair mechanisms and base substitutions (Murphy. K, 2008). If cell division proceeds before any repair events, DNA replication can reads uracil residues as thymidine residues thus generating a C to T transition. Alternatively, the uracil residue produced by the action of AID can become the substrate for the base-excision repair enzyme uracil DNA glycosylase (UNG). UNG can excise the U residue, and this process generates an abasic site in the DNA that may be replicated over by the action of error-prone polymerases (Stavnezer, 2011). These actions promote transitions and transversions at C:G pairs. Mutation at A:T pairs are considered to result from recognition of the U:G mismatch by MSH2/MSH6 which initiates a mutagenic patch repair event also involving error-prone polymerases (Di Noia and Neuberger, 2007).

Somatic hypermutation involves the introduction of point mutations throughout the variable regions of both the heavy and light chains at a high rate but occurs during only a narrow window of B cell development. Nucleotide substitutions are made at a frequency of approximately 10^{-3} per base pair per generation (Di Noia and Neuberger, 2007). All four bases can be targeted for mutation, C:G pairs and A:T pairs are targeted with approximately equal frequency and all four bases can be targeted for the introduction of mutations. However, though it is known that mutational hot spots exist, and these are often at C:G pairs found within a WRCY consensus (where W =A/T, R =A/G, and Y=C/T) (Di Noia and Neuberger, 2007).

In the case of CSR, AID mediates a deletional recombination event of the *IGH* loci and involves the C region only; where replacement of the original Cμ *IGH* chain with an

alternative C region takes place (Dunnick et al., 2009). AID catalyses CSR by deamination of several deoxycytidine residues in single stranded DNA switch (S) regions and the subsequent UNG-mediated excision of the resulting deoxyuracil residues. (Gatto and Brink, 2010; Peled et al., 2008; Stavnezer et al., 2008). Instead of filling the resulting abasic residues directly, the action of the apurinic/apyrimidinic endonucleases 1 (APE1 and APE2) serves to excise the rest of the residue consequently introducing a single or double stranded nick (the latter if the abasic sites are sufficiently close on opposite strands) in the DNA opposite the site of the original cytosine. High densities of single stranded DNA nicks in S regions flanking C region genes is suggested to produce the staggered double-stranded breaks required for CSR (Murphy. K, 2008; Stavnezer, 2011).

The progeny of the proliferating centroblasts are known as centrocytes that express surface Ig and transit to the light zone to be screened and selected for their affinity for antigen in part by signalling through their BCR. Here they receive cognate T cell help from a specialised subset of $CD4^+$ T cells known as follicular T helper cells (T_{FH}) via CD40/CD40L interactions and FDCs. T_{FH} cells produce a variety of cytokines that are thought to play role in the initiation of class switch recombination, or isotype switching, by directing AID activity to the switch recombination sequences located 5' of each unrearranged C_H region; γ, α and ε segments (Murphy. K, 2008).

The GC environment is highly dynamic, where refining of T cell-dependent humoral immunity is achieved through the positive selection of high affinity centrocytes for further amplification and is simultaneous to the induction of apoptosis in non-selected B cells. Maintenance of this environment is mediated by a variety of signals and changes in gene expression patterns. The isolation of naive $IgD^+CD27^-CD38^{lo}$ with unmutated V regions) by magnetic cell separation (MACS) and FACS-sorted centroblast ($CD38^{hi}$ $CD77^+$) and centrocyte B cell populations ($CD38^{hi}$ $CD77^-$) for microarray analyses allowed for the study

of global gene expression changes in human B cells entering a GC reaction, and it was demonstrated that the naive to CB transition is accompanied by changes in 457 different genes involved in the down-regulation of cell adhesion molecules, the acquisition of a pro-apoptotic phenotype and up-regulation of genes involved in chromatin remodeling, DNA replication and DNA repair (Klein et al., 2003). Expression of Bcl6 protein, product of the proto-oncogene *BCL6*, is restricted to the GC stage of a developing B cell (Basso and Dalla-Favera, 2012) and is required for the repression of hundreds of transcription factors during GC formation and a recent study unveiled a potential role of Bcl6 in regulating AID expression and promoting the expression of important GC functions through repression of the micro-RNA-155 (miR-155) (Basso et al., 2012). Cell fate decisions are then made as to whether the new B cell with a high affinity receptor differentiates into a memory B cell or a long lived plasma cell, or alternatively whether to re-enter the dark zone to undergo further rounds of receptor diversification by SHM. The dynamics of a human GC are depicted in Figure 1.8.

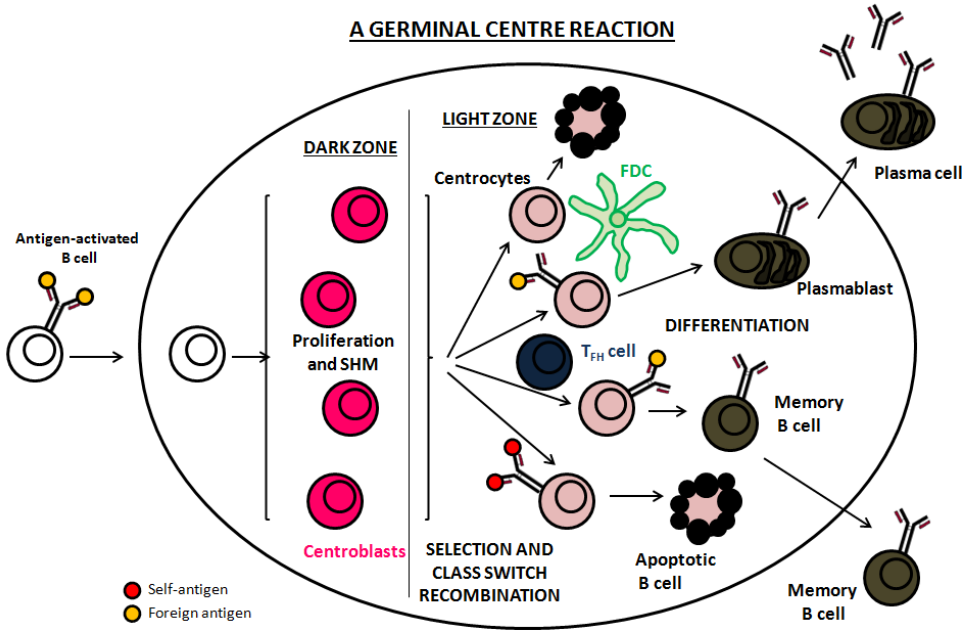


Figure 1.8 The dynamics of a human germinal centre. Antigen-activated B cells enter germinal centres where they undergo several rounds of proliferation and random somatic hypermutation in the dark zone as centroblasts. Centroblasts then migrate to the light zone, the site of selection and class switch recombination, as centrocytes where they meet one of several fates. Centrocytes lacking surface Ig or recognising self-antigen undergo apoptosis. Centrocytes expressing high affinity antibody receptors for foreign antigen are selected for differentiation into plasmablasts or memory B cells before leaving the germinal centre to continue their differentiation.

1.2 Systemic Lupus Erythematosus (SLE)

1.2.1 SLE disease outline

SLE is a chronic, incurable autoimmune disease of unknown origin with highly heterogeneous and complex clinical manifestations. This is thought to be the result of the wide range of genetic and molecular defects that contribute to SLE risk (Dorner et al.,

2011a). SLE is characterised by chronic B cell hyper-activation and the production of autoantibodies that target self-DNA and nucleoproteins (Friou, 1958; Holman and Kunkel, 1957). This process leads to the formation of pathogenic immune complexes and widespread activation of complement; the consequence is multisystem inflammation of connective tissue (Dorner et al., 2011a; Kerekov et al., 2011; Munoz et al., 2010). The main organs typically involved are the kidneys, manifesting as lupus nephritis, the skin, the eyes, the central nervous system (CNS) and the joints (Jordan et al., 2013).

The prevalence of SLE in the UK is approximately 28 per 100,000 individuals with a female to male ratio of 9:1 SLE prevalence is more common in certain countries and some ethnic groups tend to be more susceptible to SLE development. One study reported than the mean age at disease diagnosis was 6 years younger among African-American subjects and other minority groups compared with Caucasian patients (Cooper et al., 2002). The same study suggested that discoid lupus, proteinuria, anti-Sm and anti-RNP autoantibodies were more commonly seen in African-American patients though mucosal ulcers and photosensitivity were observed less frequently compared to Caucasian patients. African American and Orientals are considered to have more severe disease than SLE patients of Caucasian origin (D'Cruz et al., 2007; Lau et al., 2006). Such differences suggest a role of genetic and environmental influences on SLE development.

1.2.2 The pathogenesis of SLE

The pathogenesis of SLE remains incompletely defined, though a body of evidence strongly suggests that SLE is a syndrome resulting from aberrant interplay between a number of cell intrinsic, genetic, molecular and environmental cues.

1.2.2.1 Immune cells in SLE pathogenesis

One of the main cell intrinsic defects thought to contribute to SLE development and pathogenesis in some patients is a failure to clear apoptotic cells. This results in the uptake of previously intracellular self-constituents by macrophages that can subsequently drive an autoimmune phenotype through presentation of these antigens to B and T cells (Munoz et al., 2005; Munoz et al., 2010). Such findings have been supported by the identification of a neutrophil-specific mechanism that appears to contribute to SLE pathogenesis. Neutrophil extracellular traps (NETs) are antimicrobial web-like structures of de-condensed chromatin and cytosolic antimicrobial proteins secreted neutrophils. Several studies have provided evidence to suggest that SLE patients fail to clear NETs, typically released upon infection, due to impairment of DNase I function. This may provide a source of the self-antigen to initiate and propagate an autoimmune disease state (Hakkim et al., 2010; Knight et al., 2012). SLE is also associated with the chronic activation of plasmacytoid dendritic cells (pDCs), where it has been demonstrated that immunogenic complexes from the sera of SLE patients were composed of NET proteins and self-DNA that subsequently triggered the activation of innate pDCs via TLR9 (Lande et al., 2011). The same study also demonstrated that circulating neutrophils from SLE patients produced more NETs than healthy donors. Engagement of both the BCR and TLR9 by DNA-containing complexes activates autoreactive B cells (Leadbetter et al., 2002).

The analysis of anti-DNA antibodies and why they should be a hallmark of SLE was initially unclear since mammalian DNA itself is poorly immunogenic compared with microbial DNA (Hirabayashi et al., 2007). This was elucidated by identification that TLR9 recognises unmethylated CpG dinucleotide motifs, abundant in microbial DNA fragments. There is a low abundance of such motifs in mammalian DNA which are typically highly methylated and depleted of CpG di-nucleotides. This study demonstrated a role for CpG island-rich (50-80% GC containing DNA stretches) mammalian dsDNA fragments

approximately 600bp in length in the activation of autoreactive B cells via TLR9 (Uccellini et al., 2008).

1.2.2.1.1 T cells in SLE pathogenesis

Although this study explored the role of B cell immunoglobulin light chains in SLE pathogenesis, it is important to acknowledge the potential consequences of aberrant immunoglobulin recognition and the result of self-antigen presentation to the T cell repertoire in SLE patients. Following the initial loss of tolerance associated with SLE development, the interaction of autoreactive B and T cells in patients with the disease perpetuates and amplifies the autoimmune and inflammatory response. The resulting tissue injury further contributes to end-organ damage (Crispin et al., 2010). SLE T cells have been shown to elicit more rapid and increased early signalling events following stimulation via the TCR and CD3 signalling complex, displayed by a rapid increase in intracellular calcium concentrations and increased cytoplasmic protein tyrosine phosphorylation in freshly isolated T cells from SLE patients compared to healthy donors following anti-CD3 stimulation (Vassilopoulos et al., 1995). SLE T cells have also been shown to display aberrant regulation of apoptosis and necrosis. SLE T cells exhibit enhanced spontaneous apoptosis in vitro that correlates with disease activity and increased release of intact nucleosomes into the extracellular environment following analysis by ELISA. (Emlen et al., 1994). SLE T cells also undergo necrosis in response to oxidising agents instead of apoptosis observed in healthy subjects (Gergely et al., 2002). These abnormal processes may provide a source of self-nuclear antigen that may contribute to the propagation of autoimmunity.

T cell subsets have also been found to exhibit differences compared to healthy subjects. T cells are defined by their thymically-determined expression or absence of co-stimulatory molecules CD4 and CD8 (Crispin et al., 2010). SLE patients have increased levels of

activated $CD8^+$ T cell populations though they have been shown to be impaired in their cytotoxic ability (Stohl et al., 1999). Double negative (DN) T cells ($CD3^+CD4^-CD8^-$) have also been shown to be expanded in the peripheral blood of SLE patients. Although these cells have not directly been linked with SLE pathogenesis, they have been shown to secrete significant amounts of inflammatory IL-17 and IFN-γ, and can expand in response to CD3 stimulation in vitro (Crispin et al., 2008). IL-17^+ DN T cells were also observed in the kidneys of nephritis patients, suggesting a contribution to disease pathogenesis.

Th17 cells were identified on their ability to produce IL-17, and commitment to the Th17 lineage was demonstrated to be dependent upon expression of STAT3 and retinoic acid orphan receptor-(ROR-γt) in the presence of IL-6 and TGF-β (Alunno et al., 2012). The serum of patients with SLE has been shown to contain higher IL-17 concentrations, and the expansion of T cells producing IL-17 in peripheral blood have been identified. Th17 cell infiltration of the kidneys has also been observed in some cases of SLE (Alunno et al., 2012).

Regulatory T cell populations (T_{REGS}) constitute approximately 1-2% of peripheral blood and are an immunologically-suppressive T cell subset that have the cell surface phenotype $CD4^+CD25^+Foxp3^+$ (Cai et al., 2012). T_{REGS} may be thymically-derived (natural, nT_{REGS}) or induced peripherally through the induction of the transcriptional regulator forkhead box P3 (FoxP3). These are referred to as inducible T_{REGS} (iT_{REGS}), and it has recently been proposed that the cell surface molecule neuropilin-1 (Nrp-1) may distinguish nT_{REGS} that express high levels of Nrp-1, from iT_{REGS} (Yadav et al., 2012). It has been reported that circulating T_{REGS} are reduced in frequency in active SLE and have a reduced capacity to suppress effector T cell populations compared to healthy controls (Bonelli et al., 2008; Miyara et al., 2005). The reported imbalance of Th17 and T_{REG} cell subsets in SLE suggest that disease pathogenesis may be associated with the homeostatic control of the T_{REG}

population, though as described previously, distinguishing between cause and effect of cellular and biochemical differences implicated in SLE pathogenesis remains challenging.

1.2.2.2 Cytokines in SLE pathogenesis

Cytokines collectively regulate systemic inflammation, and it is known that several contribute to disease pathogenesis in SLE. SLE is associated with a distinctive interferon (IFN) signature, where a continuous production of IFNα correlates with severe disease in SLE patients compared to healthy controls. Type I IFNα/β is produced following infection and plays diverse roles and is involved in the innate anti-viral response, stimulating Th1 differentiation, enhancing signalling through the BCR, regulation of apoptosis and the induction of DC maturation (Braun et al., 2002; Le Bon et al., 2006; Moser et al., 2009). As previously mentioned an early event implicated in triggering SLE is the chronic activation of pDCs. This event induces the production of type I IFN that has been shown to be responsible for the unabated differentiation of monocytes into myeloid DCs that can subsequently maintain continuous activation of autoreactive B and T cells (Ronnblom and Pascual, 2008). Serum IFNα from SLE patients has been shown to induce DC differentiation from normal monocytes; a phenomenon that correlated with disease activity. It was demonstrated that the DCs induced by this method subsequently were able to present antigen captured from apoptotic cells to $CD4^+$ T cells (Blanco et al., 2001). Furthermore, IFNα has been shown to lower the activation threshold of autoreactive B cells (Le Bon et al., 2006).

In both murine and human SLE, serum levels of IL-6 are elevated. IL-6 promotes the activation and differentiation of macrophages, neutrophils and both B and T cells (Jacob and Stohl, 2011). SLE B cells spontaneously express IL6-R and produce high levels of IL-

6. It has been reported that autoreactive T cell clones promote the response by B cells and the production of autoantibodies in SLE through production of IL-6 (Takeno et al., 1997).

Tumour necrosis factor alpha (TNFα) is a cytokine capable of promoting pro- and anti-inflammatory responses depending on the environmental conditions. Evidence for a role of TNFα in human SLE pathogenesis is however contentious, with some studies reporting that elevated serum TNFα levels in patients with SLE correlates with disease activity (Studnicka-Benke et al., 1996), and others demonstrating no correlation between TNFα plasma concentrations and disease activity. The latter study even suggests a protective role for the cytokine since higher TNFα plasma concentrations were observed in patients with quiescent disease (Gomez et al., 2004).

Of further importance and subject of ongoing study is the cytokine known as B lymphocyte stimulator (BLyS), also known as the B cell–activating factor belonging to the TNF family (BAFF) that is essential for the survival of B cells. BLyS binds three different receptors: BAFFR (or BR3), B cell maturation protein (BCMA) and transmembrane activator and calcium modulator and cyclophilin ligand interactor (TACI). The receptors BCMA and TACI also bind the cytokine named a proliferation-inducing ligand (APRIL) which is closely related to BLyS (Jacob and Stohl, 2011). Important biological functions are mediated by the different interaction of these cytokines with their different receptors. Plasma cell survival is dependent upon cell surface expression of TACI and BCMA and their ligands BLyS or APRIL, whereas the maintenance of pre-PC populations is solely mediated by the abundant expression of BAFFR. Memory B cells are thought to exist independently of both BLyS and APRIL (Jacob and Stohl, 2011). Circulating levels of BLyS have been reported to be elevated in SLE patients with active disease, and high mRNA levels of BLyS and APRIL in B cells and plasma cells have been associated with levels of anti-dsDNA antibodies and disease severity (Chu et al., 2009; Petri et al., 2008).

1.2.2.3 Complement in SLE pathogenesis

It has long been recognised that the complement system plays a role in SLE. A model of SLE pathogenesis involves the activation of complement by autoantigen bound Ig immune complexes that can subsequently be deposited as inflammatory lesions in tissues (Lachmann et al., 1962; Walport, 2002). There is also an association of lower complement-mediated immune complex clearance in the sera of SLE patients with active disease compared with patients with inactive disease (Sakurai et al., 1982). Furthermore, homozygous hereditary deficiencies in several proteins of the classical pathway of complement activation are proposed to be strongly associated with the development of SLE, the most severe disease being associated with deficiencies in the C1q complex protein components and C4 protein deficiency (Pickering et al., 2000; Walport, 2002).

1.2.2.4 Environmental factors associated with SLE pathogenesis

The fact that SLE is predominantly a female disease means that sexual dimorphism forms the basis of much investigation into the origin of the disease. This has provoked a body of work investigating the involvement of estrogen in SLE, where roles have been demonstrated for estrogen in modulation of lymphoid cell growth, differentiation, proliferation, antigen presentation, cytokine production and antibody production (McMurray, 2001). There is also evidence to suggest that bacterial, viral and parasitic infections may play a role in triggering or the exacerbation of SLE (Bach, 2002). Exposure to Epstein-Barr virus (EBV) is highly associated with SLE development, where it has been described that some patients with SLE harbour EBV-containing B cells in peripheral circulation, and EBV-DNA is frequently detected in oropharyngeal secretions (James et al., 2001). However, EBV infects activated B cells and it is therefore difficult to discriminate between cause and effect. Vaccinations have also been linked with SLE development

(Aron-Maor and Shoenfeld, 2001), exemplified by the development of lupus nephritis in a female following hepatitis B vaccination (Santoro et al., 2010). Other environmental factors considered to possibly play a role in SLE development include the exposure to chemicals such as aromatic amines, hydrazines, hair dyes and tobacco smoke (Molina and Shoenfeld, 2005).

1.2.3 The genetic basis of SLE

Genetic background can influence the risk of an individual developing SLE, illustrated from studies involving monozygotic and dizygotic twins and other siblings which exhibit concordance rates of greater of 24% and 2-5% respectively (Deapen et al., 1992). Since 2008, numerous genome wide association studies (GWAS) have been performed in patients with SLE in various ethnic populations and identified and confirmed the association of 40 genes/risk loci with SLE development (Cui et al., 2013).

Numerous genetic polymorphisms associated with human SLE that relate to immune cell signalling (Budarf et al., 2011). In 2008, Kozyrev and colleagues applied a genome-wide scan of 85,042 single nucleotide polymorphisms (SNPs) and identified an SLE-associated nonsynonymous substitution (rs10516487, R61H) in the gene encoding the B-cell adaptor protein with ankyrin repeats (BANK1) Susceptibility to SLE in European and Asian populations has been associated with polymorphisms of the *BANK1* gene in both functional domains and regulatory sites of the protein (Guo et al., 2009; Kozyrev et al., 2008). BANK1 consists of 13 tyrosine residues susceptible to phosphorylation. BANK1 is bound by LYN and SYK and is required for the optimal induction of Ca^{2+} mobilisation by LYN-mediated phosphorylation of type I and II IP3 receptors. Although *BANK1* knock-out mice display no autoimmune phenotype, BANK1 has also be shown to interact with BLK, also an SLE susceptibility allele encoding a tyrosine kinase, both epistatically and at the protein level

(Castillejo-Lopez et al., 2012). BANK1 and BLK bind upon BCR stimulation *in vitro*, and together these data suggest a role for these proteins in a potentially biologically relevant cell signalling pathway in a disease where hyper B cell activation is a defining feature. Such pathways may have a role in the regulation of gene rearrangements events that ensure that *IGKV* segments with undesirable features are not expressed.

With respect to T lymphocyte genetic risk loci for SLE, there is much interest in the protein tyrosine phosphatase, non-receptor type 22 (*PTPN22)*, where identification of a functional R620W SNP in exon 14 of *PTPN22* was first described in type I diabetes and later identified in SLE and RA (Begovich et al., 2004; Bottini et al., 2004; Chung and Criswell, 2007; Kyogoku et al., 2004). Exon 14 encodes a lymphoid protein tyrosine phosphatase (LYP) and the risk polymorphism R620W results in an arginine to tryptophan substitution. The resulting gain-of-function variant cannot interact with the negative regulatory kinase CSK resulting in unabated intrinsic phosphatase activity. Ultimately T cell activation cannot be suppressed. It has also been described that a loss-of-function variant of PTPN22, R263Q, conveys a missense substitution within the catalytic domain of LYP that also alters its intrinsic phosphatase activity and is also associated with SLE risk (Orru et al., 2009).

The programmed cell death 1 gene (*PDCD1*) encodes the immunoinhibitory receptor PD1 and intronic SNPs within this gene have also been associated with autoimmunity. A SNP located in the fourth intron of the gene has been shown to confer an altered binding site for the runt-related transcription factor 1 (encoded by the RUNX1 gen) which is a transcription factor involved in HSC differentiation (Prokunina et al., 2002).

Consistent with the IFN signature associated with SLE, the up-regulation of IFN-responsive genes has also been observed in gene profiling studies of SLE patients (Baechler et al., 2004; Crow and Kirou, 2004). Interferon regulatory factor 5 (IRF5) is also associated with SLE susceptibility, with an odds ratio (OR) of greater than 1.5. Three functional

variants of IRF5 have been described, each having different effects on interferon signalling and conferring distinct levels of risk to SLE (Cunninghame Graham et al., 2007; Graham et al., 2007).

Interleukin-1 receptor-associated kinase (IRAK1) is a serine/threonine kinase involved in Toll/IL1R (TIR) signalling and has been implicated as an X-chromosome encoded risk factor for SLE. Five SNPs spanning the IRAK1 gene were identified in a study of over 5,000 SLE patients and healthy controls and demonstrated that these SNPs displayed disease association in adult- and childhood-onset SLE and across four different ethnic groups (Jacob et al., 2009).

The signal transducer and activator of transcription 4 gene is located on chromosome 2q and encodes the STAT4 transcription factor. STAT4 is typically a latent cytosolic protein but is activated by several cytokines including type I IFN, IL-12 and IL-23 that are predominantly produced by DCs. STAT4 stimulates the transcription of several genes involved in regulation differentiation of Th1 and Th17 cell lineages (Nishikomori et al., 2002; Remmers et al., 2007; Watford et al., 2004). A SNP haplotype in the third intron of STAT4 is associated with susceptibility to SLE and RA (Remmers et al., 2007). These observations were confirmed in another study that identified that this SNP haplotype was associated with a more severe SLE phenotype where nephritis was prevalent (Taylor et al., 2008). STAT4 risk variants have now been identified in African-American, European, Hispanic-American, Asian and Latin-American populations and interestingly, an additive effect has been described between STAT4 and IRF5 for increasing predisposition to SLE (Abelson et al., 2009; Namjou et al., 2009).

An SLE risk haplotype has also been identified in the tumour necrosis factor-α-induced protein 3 (*TNFAIP3*) gene and studies in the context of several autoimmune diseases have demonstrated an association with Crohn's disease, RA and psoriasis (Moser et al., 2009).

TNFAIP3 encodes a ubiquitin-editing and zinc finger protein named A20. This protein suppresses pro-inflammatory responses mediated via NFκB, downstream stimuli induced through TLRs, TNFR and IL1R (Boone et al., 2004). *Tnfaip3*$^{-/-}$ mice exhibit severe spontaneous multisystem inflammation and are hyper-responsive to LPS and TNF and fail to inhibit TNF-induced NF-κB responses (Lee et al., 2000). A recent study involved the generation of chimeric mice by crossing *Cd19-Cre* knock-in mice with mice carrying loxP flanked *Tnfaip3* alleles to selectively abolish intrinsic expression of A20 in CD19^{+} B cells. These *Tnfaip3*$^{fl/fl}$ *CD19-Cre* mice were largely healthy but exhibited reduced CD40 and BCR-triggered NF-κB signals and increased generation of autoantibody producing cells (Tavares et al., 2010).

1.3 B cells in SLE

Large bodies of evidence implicate virtually all immune cell types in the pathogenesis of SLE. This undoubtedly reflects the complexity of the condition and difficulty in understanding if differences in cell phenotype and behaviour are a cause or consequence of the disease. What is acknowledged however, is that SLE involves a profound breakdown in B cell tolerance accompanied by chronic B cell hyperactivity and the production of autoantibodies, where multiple factors thought to be involved in disease pathogenesis and are not mutually exclusive and are unique to each patient.

1.3.1 B cell tolerance checkpoints in SLE

Two discrete checkpoints in early B cell tolerance have been shown to be responsible for the removal of autoreactive B cells in healthy subjects. These checkpoints were overcome in three untreated adolescent patients with SLE (Wardemann et al., 2003; Yurasov et al.,

2005). This study involved the isolation of single early emigrant B cells ($CD19^+CD10^+IgM^+CD27^-$) and mature naive B cells ($CD19^+CD10^-IgM^+CD27^-$) from three SLE patients the subsequent production of cDNA libraries. This allowed for the cloning and expression of a total of 222 SLE-antibodies and subsequent screening for their autoreactivity by HEp-2 cell lysate ELISA. Compared to healthy controls where HEp-2 reactive antibodies were detected in 40.7% and 20.4% of new emigrant and mature naive B cells respectively, SLE HEp-2 reactive antibodies similarly accounted for 40.5-59.4% new emigrant B cells, however 41.5-50% of the antibodies produced by mature naive SLE B cells bound HEp-2. These data suggest a failure of a checkpoint between early emigrant and mature B cells in SLE (Yurasov et al., 2005). The following year, the same group examined the same B cell tolerance checkpoint in six SLE patients in clinical remission through the cloning of 278 antibodies from mature naïve B cells and subsequent analysis of their HEp-2 cell lysate binding capacity. Despite individual variability, mature naïve B cells from SLE patients in remission retained increased autoreactivity compared to controls, though the HEp-2 reactivity observed was lower than that of patients with active disease (Yurasov et al., 2006).

A tolerance checkpoint in late B cell development has also been described. It has been shown that the spleen and BM of non-autoimmune, healthy mice carry high frequencies of ribonucleoprotein Smith (Sm)-specific pre-plasma cells (pre-PCs). Autoantibodies to Sm antigen are commonly detected in human SLE patients and have been associated with a higher risk of renal involvement (Alba et al., 2003). These Sm-specific pre-PCs were found to be predominantly located at the T/B cell borders and had phenotypic characteristics of PC differentiation, though the Sm-specific pre-PCs were not autoantibody secreting and had a higher frequency of cell death compared to other non-Sm binding pre-PCs (Culton et al., 2006). This apparent regulation of autoimmunity took place prior to detection of B lymphocyte-induced maturation protein-1 (Blimp-1) expression, which drives terminal

differentiation of B cells into PCs (Shaffer et al., 2002). These findings suggest the existence of a tolerance checkpoint before the terminal differentiation of pre-PCs to mature PCs. Interestingly, this peripheral checkpoint was found to be deficient in lupus-prone MRL/lpr mice (Culton et al., 2006).

1.3.2 B cell phenotypes in SLE

The main characteristics of disturbed B cell homeostasis in SLE is the expansions in $CD24^{hi}CD38^{hi}$ transitional, switched-memory and $CD27^{++}$ plasma cell populations in peripheral blood and an immature B cell repertoire that is predominantly self-reactive (Dorner et al., 2011b; Jacobi et al., 2008; Jacobi et al., 2009; Odendahl et al., 2000; Wardemann et al., 2003). Features of SLE B cell subsets will be described in greater detail in the following sections.

1.3.2.1 Transitional B cells in SLE

Transitional B cells are present at a frequency of approximately 2-3% in healthy peripheral blood, though they are increased to approximately 6-7% as a proportion of whole peripheral blood in SLE patients (Sims et al., 2005). As described in section 1.1.8.2, human T2 transitional B cells have been shown to be selectively recruited to GALT where they acquire an activated phenotype, potentially representing a novel specificity checkpoint (Vossenkamper, 2013). Importantly, since SLE is associated with aberrant tolerance checkpoints, this study went on to examine the path of T2 transitional B cells in SLE. They illustrated that SLE lymphocytes expressed significantly lower levels of $\alpha4\beta7$, implying impaired ability to home to gut tissue, and a failure to up-regulate this integrin on SLE blood B cells compared to healthy B cells following *in vitro* stimulation with CpG and

retinoic acid. These findings suggest that a failure to access GALT may represent an additional aberrant tolerance checkpoint associated with SLE, though it remains unclear how these observations relate to individual disease phenotypes.

1.3.2.2 Naïve, memory and plasmablast subsets in SLE

Patients with SLE tend to have marked B lymphocytopenia that affects the naïve $CD19^{+}CD27^{-}$ B cell pool more than the $CD19^{+}CD27^{+}$ memory B cell populations. This is thought to be the source of the increased proportion of $CD27^{+}$ peripheral B cells observed in SLE rather than expansion of the $CD27^{+}$ subset itself (Odendahl et al., 2000). Despite increased frequencies of $CD27^{++}$ peripheral B cells detected in both active and inactive SLE patients compared to healthy donors, the absolute numbers of $CD27^{++}$ plasma cells ($CD38^{+}CD19^{dim}sIg^{lo}CD20^{-}CD138^{+}$) were increased only in the peripheral blood of SLE patients with active disease. The $CD27^{++}$ cell population was rarely observed in the peripheral blood of healthy control subjects. In addition, following immunosuppressive therapy, it is the $CD27^{++}$ plasma cells and naïve $CD27^{-}$ B cell pools that become markedly decreased in peripheral blood compared to similar patients that were untreated (Odendahl et al., 2000). Furthermore, the $CD27^{++}$ B cell subset significantly correlates with disease activity and anti-dsDNA antibody titres (Jacobi et al., 2003)

It has been reported that $IgD^{-}CD27^{-}$ B cells are a heterogeneous memory B cell population expressing IgM, IgG or IgA that are associated with increased renal disease, the presence of autoantibodies and disease activity in SLE patients (Wei et al., 2007). The nature of these $IgD^{-}CD27^{-}$ memory B cells has been explored in SLE and it has been suggested that CD95 may help to identify them as a homogenous memory B cell population associated with SLE since the $IgD^{-}CD27^{-}CD95^{+}$ cell surface phenotype excludes $CD10^{+}$ transitional B cells

(Jacobi et al., 2008). They too observe that this memory B cell subset is increased in SLE patients with disease flares.

1.3.2.3 Regulatory B cells in SLE

Our increasing understanding for the roles of B cells with regulatory capacity in the control of autoimmunity make them highly interesting therapeutic targets, where enrichment of these populations or enhancement of their suppressive capacity is a current therapeutic goal (Blair et al., 2009). Regulatory B cells have been implicated in the amelioration of several autoimmune conditions in both mouse models and human subjects. The T2-MZP B cells subset previously mentioned is thought to contain IL-10-producing B regulatory cells, and it has been shown the transfer of in vitro anti-CD40 generated T2 B cells (T2-like-Bregs) improve renal disease and survival in lupus prone MRL/lpr mice by a mechanism dependent on IL-10 (Blair et al., 2009). CIA chimeric mice lacking IL-10-producing B cells have also been shown to exhibit exacerbated arthritis compared to WT controls, and comparably increased populations of T cells with Th1 and Th17 phenotypes, suggesting that the IL-10 producing B cells can play a prominent immunoregulatory role in an inflammatory environment (Carter et al., 2012).

Studies using human $CD19^{+}CD24^{hi}CD38^{hi}$ regulatory B cells isolated from SLE patients have demonstrate that this subset tend to produce less IL-10, be refractory to CD40 stimulation and have an impaired suppressive capacity compared to healthy equivalents (Blair et al., 2010). Human regulatory B cells from healthy subjects have also been shown to hold the capacity to suppress $CD4^{+}CD25^{-}$ T cell proliferation and their secretion of TNF-α and IFN-γ by conversion into functionally suppressive T_{REGS} cells, in part due to the action of IL-10 (Flores-Borja et al., 2013). The similar $CD19^{+}CD24^{hi}CD38^{hi}$ B cell subset

isolated from RA patients failed to generate functional T_{REGS} from CD4+CD25 T cells, though this has yet to be demonstrated in SLE.

1.3.3 Immunoglobulin light chains in SLE

Both at the genetic level and level of expression, the human immunoglobulin light chain repertoire displays preferential usage of some *IGKV and IGKL* gene segments. In light of this observation, several studies have investigated the contribution of molecular events that underlie the expression of protein Ig during B cell repertoire development (Dorner et al., 1998; Foster et al., 1997). The *IGK* light chain repertoire of healthy individuals was examined during an analysis of more than 350 productive and 250 non-productively rearranged *IGK* gene sequences that were cloned and sequenced from single FACS-sorted cells. Non-productive rearrangements of *IGK* are not transcribed and translated into protein so studying them reflects the inherent VκJκ rearrangements during the development of a B cell, and as a result they have no influence over the fate of the B cell (Foster et al., 1997). This study observed the preferential usage of six IGKV gene segments; *IGKV3-20* (A27), *IGKV3-15* (L2), *IGKV3-11* (L6), *IGKV2-30* (A17), *IGKV1-5* (L12a) and *IGKV1-39/1D-39* (O12/O2). $J_{\kappa}1$ and $J_{\kappa}2$ usage was also overrepresented and accompanied by an underrepresentation of $J_{\kappa}3$ and $J_{\kappa}5$ within non-productive rearrangements, suggestive of a molecular decision for this bias. A consistent CDR3 length was observed in both productive and non-productive rearrangements of *IGK* genes (Foster et al., 1997). 23% of productive VκJκ rearrangements carried somatic hypermutations, in contrast to the non-productively rearranged VκJκ sequences where just 7% of these were mutated, reflecting functional activation by the KDE (Foster et al., 1997). Several of these observations were also shown in a similar study of the same year that analysed 103 rearranged sequences of *IGK* following cloning from the cDNA isolated from healthy PBMCs where 65% of the

sequences involved usage of *IGKV3-20* (A27), *IGKV3-11* (L6), *IGKV3-15* (L2) and *IGKV2D-28* and/or *IGKV2-28* (Juul et al., 1997).

A study investigating the contribution of receptor editing to SLE pathogenesis analysed the *IGKV* gene segments usage by $CD19^+$ B cells isolated from an untreated male SLE patient. They observed no difference in either *IGKV* or *IGKJ* gene segment usage in rearrangements that were non-productive compared to equivalent rearrangements of a healthy adults (Dorner et al., 1998). However, over-representation of *IGKV1D-16, IGKV1-37/1D37 and IGKV4-1* was observed in the productively rearranged *IGKV* repertoire of this patient compared to healthy equivalent rearrangements. Interestingly, over-representation of *IGKV4-1* was observed at the level of expression only, since this bias was not observed in the non-productive *IGKV* rearrangements, suggesting that the inherent Ig gene rearrangement process robust in this SLE patient. Increased usage of the *IGKJ5* segment compared to healthy controls was also found in the expressed kappa light chain repertoire of this patient, suggesting more apparent receptor editing of light chains in this individual with SLE.

As described in section 1.3.1, a study of B cell tolerance checkpoints in the context of SLE involved the cloning of antibodies from cDNA libraries produced from the single cells isolated from three untreated SLE patients. Alongside examination of self-reactivity at different stages of B cell development, Ig gene usage was also studied (Yurasov et al., 2005). As observed previously, an over-representation of the gene segment *IGKV4-1* was observed in the expressed repertoire of newly emigrated and mature naive B cells from two of the three SLE patients. A study the following year examined the frequency of somatic mutations in groups of productive and non-productive *IGK* rearrangements sequenced from $CD19^+$ B cells from an SLE patient. Compared to healthy donors, a statistically significant increase in mutational load was observed in the non-productively rearranged repertoire

(Dörner et al., 1999). This group of sequences theoretically should be unmutated since the inactivating rearrangement of the KDE deletes the *IGKC* segment and the enhancers flanking it (Das et al., 2009). Significantly increased frequencies of mutations were also observed within the productively rearranged sequences, suggesting that enhanced mutational activity may be a feature of this SLE patient's disease. Interestingly, the frequency of mutations within the non-productive *IGK* rearrangements in this SLE patients was greater than those observed in the groups of productive rearrangements (Dörner et al., 1999).

With respect the human *IGL* light chain repertoire, single cell PCR studies have described the over-expression and under-representation of specific *IGL* gene segments in groups of productively and non-productively rearranged sequences (Farner et al., 1999). These data provide further evidence for the existence of positive and negative selection mechanism underlying development of the Ig repertoire. In addition, enhanced mutational activity has been observed within productive and non-productively rearranged sequences isolated from the $CD19^+$ $Ig\lambda^+$ B cells of an SLE patient, suggesting globally aberrant light chain mutational activity in SLE (Dorner et al., 2001).

1.3.4 Diagnosis of SLE

There is no single diagnostic test for SLE, though the hallmark of the disease is the presence of circulating autoantibodies targeted against an array of intracellular antigens. Examples include anti-nuclear antibodies (ANAs), anti-dsDNA, anti-Ro, anti-Sm, anti-RNP, anti-La which are frequently detected years prior to the onset of the clinical pathologies associated with the disease (Arbuckle et al., 2001). New diagnoses are made based on clinical

assessment and fulfillment of just 4 of 11 classification criteria established by the American College of Rheumatology (Criswell, 2008; Hochberg, 1997; Tan et al., 1982).

Since no single biomarker adequately marks disease activity. SLE disease activity is assessed by a number of clinical indices. Two commonly used disease activity score indices used are the Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) and British Isles Lupus Assessment Group (BILAG) index (Yee et al., 2008). These scores have since been updated and improved to the (SLEDAI)-2000 index and the (BILAG)-2004 index following revisions. The SLEDAI scoring system of 24 items and is used as a global measure of SLE disease activity, 8 of which are represented by laboratory results such as blood counts, complement component levels and autoantibody titres. The remaining 16 include clinical manifestations such as arthritis, mouth sores, muscle weakness and neurological disturbances. All scores are assigned based on whether these features are present or absent in the previous 10 days prior to consultation (Gladman et al., 2002).

The BILAG index attempts to capture varying degrees of change in severity of different organs or systems simultaneously involved (Isenberg, 2007). Clinical features that constitute the most active form of disease in an organ or system are designated the A (for action) clinical score. Groups of clinical features that reflect ongoing but reduced levels of disease activity are assigned a clinical score of B (for Beware); a clinical score of C (for Contentment) aims to acknowledge mild disease that requires symptomatic treatment only and a clinical score of D (for Discount) marks individual organs and systems that were once, but are no longer active. An E (for no Evidence) is assigned to organs or systems that have never been a feature of the patients' disease (Isenberg, 2007). This scoring system was designed to provide a testable hypothesis of whether a patient should be treated with larger doses of immunosuppressive drugs. A recent assessment compared the reliability of both the SLEDAI-2000 and BILAG-2004 indices to conclude that the BILAG-2004 index

was more able to indicate a requirement for increased treatment in patients with active SLE (Yee et al., 2008).

1.3.5 B cell targeted therapies in SLE

The range of therapies used to treat SLE are largely immunosuppressive and the consequence of the clinician's judgment of individual patients. Immunosuppressive treatments include powerful cytotoxic drugs such as azathioprine, mycophenolate mofetil and cyclophosphamide, though these regimes are non-specific and pose significant risk of infection (Yong and D'Cruz, 2008). The anti-malarial hydroxychloroquine (HCQ) is also a commonly administered SLE treatment. A recent study has demonstrated that early treatment of patients with HCQ suffered from less disease-associated damage three years after diagnosis (Akhavan et al., 2013). These treatments do not target specific pathways relating to SLE pathology. A large focus of study into new therapies for SLE patients predominantly aims at targeting B cells due to their wide range of effector functions, being not only responsible for the production of autoantibodies, but presenting antigen to other immune cells and producing a cocktail of cytokines which are immunoregulatory or immunogenic in certain environments (Townsend et al., 2010).

Rituximab (RTX) is a chimeric monoclonal anti-CD20 antibody that has become a mainstay in the therapy of a broad variety of B-cell malignancies. Over the last decade it has become more frequently administered to patients with a variety of autoimmune disorders, including use in the treatment of SLE patients with severe disease who are refractory to immunosuppressive therapy. Although the mechanism of B cell depletion by RTX is poorly defined, it is thought to induce complement-dependent cytotoxicity following recognition of rituximab-coated $CD20^{+}$ B cells (B cell precursors from the pre-

B cell stage onwards and mature B cells)(Vossenkamper et al., 2012). Rituximab treatment is also thought to initiate antibody-dependent cell-mediated cytotoxicity (ADCC) mediated by NK cells (Weiner, 2010).

The effectiveness of Rituximab has been investigated in several open-label prospective and retrospective studies. A review of studies involving 188 rituximab treated SLE patients identified that 171 (91%) demonstrated a significant improvement in one or more of their SLE clinical manifestations following treatment. However, 46 of these patients relapsed, despite 80% of those relapsing having demonstrated a clinical response to the initial treatment (Ramos-Casals et al., 2009). Furthermore, Phase III randomised controlled trials, the EXPLORER (non-renal SLE) and LUNAR (lupus nephritis SLE) trials did not demonstrate superior outcome of rituximab treatment over standard therapies (Merrill et al., 2010; Vossenkamper et al., 2012). B cell depletion was however associated with statistically significant improvement in several serological markers of disease activity such as dsDNA titres and complement C3 levels (Vossenkamper et al., 2012). The re-population of peripheral B cells following rituximab treatment has been reported to take place between 6 and 10 months following depletion and bears similarities with the B cell repopulation observed following BM transplantation (Anolik et al., 2007; Bemark et al., 2012). During reconstitution of the B cell pool in lymphoma patients following RTX-induced B cell depletion, the majority of cells have an immature and transitional phenotype (Abulayha et al., 2010; Anolik et al., 2007).

Belimumab, the first new drug to be approved for the treatment of SLE in 50 years, and represents the first biological agent to be licensed for the treatment of auto-antibody positive adult SLE (Lutalo and D'Cruz, 2012). Belimumab is a fully humanized monoclonal antagonist antibody that targets and inhibits the action of the B cell survival factor BLyS, or BAFF. BAFF is found in high concentrations in the serum of SLE patients (Liu and

Davidson, 2011; Pers et al., 2005). (Vossenkamper et al., 2012). A study investigating the long-term effects of BLyS inhibition monitored 17 patients undergoing Belimumab treatment alongside standard treatment regimen. Between days 84 and 168 B cell frequencies fell to below baseline levels and flow cytometry analysis revealed that this was a result of a decrease in naive and transitional B cells, with $CD27^{+}IgD^{+}$ memory cells and plasmablasts decreasing only after day 532. $CD27^{+}$ IgD^{-} cells and T cell subsets were found to be unchanged during the course of treatment (Jacobi et al., 2010).

1.3.6 Serum free light chains

The clonal overproduction of immunoglobulin light chains is associated with a variety of hematological malignancies. These light chains circulate without covalent association to Ig heavy chains and can readily be detected in peripheral blood and urine (Davids et al., 2010). It is thought that monitoring FLCs is a useful measure of disease activity due to the short half-lives of Igκ and Igλ light chains in blood; 2-4 and 3-6 hours respectively, compared with their half-life of approximately 21 days when in intact Ig molecules (Davids et al., 2010). A recent study compared serum free light chain (FLC) levels in 75 patients with SLE and the same number of patients with RA, and also compared these against other biomarkers of SLE, such as dsDNA Ab titres. They reported that FLC was higher in SLE than RA; both were higher than FLC levels of healthy controls. Total FLC showed moderate to strong correlation with SLEDAI, proposing the use of serum FLC analysis as a biomarker for SLE disease activity (Aggarwal et al., 2011).

1.4 Study aims

This thesis seeks to explore the human immunoglobulin kappa light chain repertoire in health and the context of the chronic autoimmune disease SLE.

Chapter 3 examines the origin of the light chain biases previously observed in SLE by a high-throughput sequencing approach. This investigation was designed to assess the inherent usage of *IGKV* and *IGKJ* gene segments by the mature naïve B cell repertoire of three healthy individuals and three SLE patients. These data were gathered from amplification of *IGK* rearrangements from gDNA of sorted B cell populations. Rearranged *IGK* transcripts were also amplified from the cDNA of populations of sorted B cells in order to examine *IGKV* and *IGKJ* gene usage within the expressed light chain repertoire both in health and SLE.

Chapter 4 investigates the rearrangement of the kappa deleting element, a mechanism that functionally inactivates rearranged kappa light chains that are in the incorrect genetic reading frame, or encode unwanted specificities. The frequency of KDE rearrangements in health and SLE have been analysed, alongside examination of the nature of KDE rearrangement junctions.

Chapter 5 investigates the cell surface expression of immunoglobulin light chains by healthy and SLE $CD19^{+}$ B cells, and brings together the data presented in chapters 3 and 4 to discuss the implications of these findings in the pathogenesis of SLE.

Chapter 2
Materials and Methods

2.1 SLE patients and peripheral blood samples

SLE patients attending clinic at the Louise Coote Lupus Unit at St. Thomas' Hospital were invited to participate in the study following written informed consent from the KCL Infectious Diseases BioBank working under the authority of the Southampton and South West Hampshire Research Ethics Committee (REC Approval: REC09/H0504/39 – The Control of Inflammation in Immunity and Autoimmunity; IDB ref JS-1b). All SLE patients collected fulfilled the American College of Rheumatology classification criteria, with at least four of the 11 criteria are present, serially or simultaneously (Hochberg, 1997). Information on organ involvement, auto-antibody titres and disease duration were also recorded, alongside any treatments each patient had once been prescribed. The clinical information of each patient studied can be found in Appendix 2. All SLE patients studied were female. Peripheral blood samples were collected in vacutainers (BD) containing K_2EDTA. The buffy coats of all healthy subjects studied were purchased from the Blood Transfusion Service, Tooting, UK.

2.1.1 PBMC isolation

Fresh blood samples were diluted 1:1 with RPMI medium + 10% FCS prior to layering approximately 35 ml over 15 ml Ficoll-Hypaque solution in a 50 ml Falcon tube. Tubes were centrifuged for 25 mins at 400g with slow acceleration and without brakes. Peripheral blood mononuclear cells (PBMCs) were removed from the Ficoll/plasma interface and washed in ~4°C RPMI medium + 10% FCS for 5 mins at 1200 rpm. The supernatant was removed and cell yield was approximated by means of a hemocytometer. Cells not to be used immediately were re-suspended in FCS supplemented with 10% DMSO (freezing medium) at an appropriate volume for storage in 2 ml cryovials at an approximate frequency of $5x10^6$ cells/ml.

2.2 Working Solutions

Autoclaved phosphate buffered saline (PBS) solution was used throughout this study at a working dilution of 1X, its applications including immunostaining of PBMCs, cell sorting and DNA isolation.

10X PBS: 80 g NaCl, 2 g KCl, 14.4 g Na_2HPO_4 and 2.4 g KH_2PO_4 (all Sigma, UK) were dissolved in 800 ml dH_20 and the pH adjusted to 7.4 with concentrated HCl. Before autoclaving, the final volume was increased to 1 L with dH_20.

2.2.1 Solutions for High-Throughput Sequencing

2.2.1.1 Sort-Lysis RT Buffer (SLyRT) Buffer: 180 µl of this buffer was used to re-suspend B cell populations following cell sorting. For single cells sorting, individual cells were sorted into 96-well plates containing 18µl SLyRT buffer per well.

Table 2.1 Recipes and concentrations for preparation of SLyRT buffer

Reagents	Initial concentration	1 reaction volumes (µl)	50 reaction volumes (µl)	Final reaction concentration
First-Strand RT buffer (Invitrogen)	5X	8	400	1X
pd(N)6 (QIAGEN)	50 ng/µl	12	600	15 ng/µl
Triton X-100 (Sigma)	5%	1	50	0.13% (v/v)
RiboSafe RNase inhibitor (40U/ml, Bioline)	40 U/µl	2.5	125	2.5 U/µl
DTT (Invitrogen)	0.1 M	4.5	225	11.25 mM
dNTP mix (Promega)	10 mM each	2	100	500 µM
PCR-graded H_2O	NA	6	300	NA

2.2.2 Solutions for Agarose Gel Electrophoresis

2.2.2.1 50X Tris/Acetate/EDTA (TAE) buffer: 242 g TRIS Base, 57.1 ml glacial acetic acid, 100ml 0.5M EDTA (all Sigma, UK) were dissolved in 800 ml of dH_20 and made up to 1 L following complete dissolution. Final 1X working concentration: 0.04M Tris-acetate, 0.001M EDTA.

2.2.2.2 10X Tris/Borate/EDTA (TBE) buffer: 108 g Trizma base, 55 g boric acid and 8.3 g EDTA (all Sigma, UK) were dissolved in 800 ml of dH_20 and made up to 1 L following complete dissolution. 10X TBE stock buffer was diluted 1:10 to make up a 1X working TBE solution.

2.2.2.3 Ethidium Bromide stock (0.5 μl/ml): 10 mg ethidium bromide was dissolved in 1 ml dH_20 was diluted 1:20 with 1X TBE to make up a working solution of 50 ng/100 ml.

2.2.2.4 Gel loading buffer: 40 g sucrose, 0.25 g bromphenol blue and 0.25 g xylene cyanole were dissolved in 100 ml ddH_20.

2.2.3 Solutions for Flow Cytometry and cell culture studies

2.2.3.1 RPMI-1640 medium: RPMI-1640 medium containing 1% L-glutamine (Gibco) was supplemented with 1% PenStrep (Sigma), 1% Amphotericin B and 10% heat inactivated FCS at 56 °C for 30 minutes.

2.2.3.2 FACS buffer: 2 % Fetal Calf Serum (FCS) (seraLab) in 1X PBS.

2.2.3.3 Glycine wash buffer pH 3.0: 0.2M glycine solution was adjusted to pH3 with acetic acid

2.2.3.4 RPMI-1640 acidic wash buffer pH 2.5: 1640 medium containing 1 % L-glutamine (Gibco) was adjusted to pH 2.5 with acetic acid.

2.2.2.5 Phorbol myristate acetate (PMA): PMA was used to stimulate B cells in co-culture with allogeneic $CD3^{+}CD4^{+}$ T cells at a concentration of 50ng/ml.

2.2.4 Solutions for Cloning

2.2.4.1 LB agar: 1g Tryptone (DIFCO), 0.5g Yeast Extract (DIFCO), 0.5g NaCl (Sigma) and 1.5g Bacto agar (BD) were dissolved in 100ml dIH_20 and the solution autoclaved at 120°C for 20 minutes to sterilise. The agar was allowed to cool for 15 minutes before adding 200 µl 0.5 µl/ml ampicillin (Sigma), 200 µl 0.5 µl/ml X-Gal (Promega) and 0.5 ml IPTG. The solution was swirled to mix the reagents and subsequently 20 ml was pipetted into petri dishes prior to blue/white screening of cloned PCR products. Prior to sending off individual clones for sequencing, the same LB agar was produced though only 200 µl 0.5 µl/ml ampicillin was added to the autoclaved LB and 200 µl added to each well of a 96-well plate.

2.2.4.2 SOC medium: 2.0g Tryptone (DIFCO), 0.5g Yeast Extract (DIFCO), 1 ml 1M NaCl (58.44g/L) and 0.25 ml 1M KCl (74.55g/L) were dissolved in 97 ml dH_20 and the solution was autoclaved at 120°C for 20 minutes to sterilise before allowing to cool. 1ml 2M Mg^{2+} stock (1M $MgCl_26H_20$/1M $MgSO_47H_20$, filter-sterilised and 1ml 2M sucrose) was added and made up to 100 ml with sterile dIH_20.

2.2.4.3 Isopropylthio-β-galactoside (IPTG) stock (0.1M): 0.6 g IPTG (Promega) was added to 20ml dIH20 and the solution was sterilised by filtering through a 0.22μM filter. Aliquots were stored at -20°C.

2.2.4.4 Ampicillin stock (50mg/ml): 1g ampicillin (Sigma) was added to 20ml dIH_20 and the solution was sterilised by filtering through a 0. 22μM filter. Aliquots were stored at -20°C.

2.3 Preparation of high-throughput sequencing (HTS) samples

2.3.1 B-cell isolation and cell sorting

Peripheral blood mononuclear cells (PBMCs) were isolated from 3 anonymous healthy single donor buffy coats obtained from the blood transfusion service and 3 SLE patients using Ficoll-Paque Plus (GE Healthcare). Samples from patients with SLE were collected through the King's College London Infectious Diseases Biobank (IDB) and written consent was obtained in accordance with Guy's Hospital research ethics committee. The SLE patients were females aged 40 (Afro-Caribbean, 12 years history of SLE, ANA+, treated with mycophenolate mofetil and hydroxychloroquine), 59 (Caucasian, 4 year history of SLE, ANA+, treated with Leflunomide) and 32 (Caucasian, 3 year history of SLE, ANA+, treated with prednisolone and hydroxychloroquine).

PBMCs were stained with CD19-PerCP-Cy5.5, Igκ light chain-APC-H7 (BD Biosciences PharMingen) and Igλ light chain-Pacific Blue (Biolegend) and LIVE/DEAD Fixable Aqua (Life Technologies) for 15 mins at 4°C before cell sorting on the FACSAria machine (BD Biosciences PharMingen.) $CD19^+$, single $Ig\kappa^+$ and $Ig\lambda^+$ expressing B cells were collected

separately into 180 μl of Sort-Lysis RT buffer (SLyRT). Unlike healthy control samples, SLE cases did not always show clear segregation of Igκ- and Igλ-expressing cells due to apparent light chain inclusion in some cases (Fraser *et al*, unpublished). The cases used in this study were those where discrete Igκ- and Igλ-expressing populations could be isolated.

2.3.2 HTS primer sequences for Vκ-Jκ (gDNA) and Vκ-Cκ (cDNA) amplification from sorted B cell populations

HTS PCR amplification of gDNA and cDNA was performed in two rounds, where the same primer sequences were used in both reactions, though the second round PCR primers were MID-tagged, for subsequent identification of cell subsets and the individuals they were derived from. Designation of MID primers is described in detail in Chapter 3.

Table 2.2 Sequences of PCR primers used for first round amplification of *IGK* rearrangements for the high-throughput sequencing analysis.

PCR1 primer name	Primer sequence
VK1	CATCCAGWTGACCCAGTCTCC
VK2	GATATTGTGATGACCCAGWCT
VK3	GACRCAGTCTCCAGCCACCCTG
VK4	GACATCGTGATGACCCAGTCT
VK5	GAAACGACACTCACGCAGTCT
VK6	GAAATTGTGCTGACTCAGTCT
JK1/4	TCCACCTTGGTCCCTYSGCCG
JK2	CTCCAGCTTGGTCCCCTGGCCA
JK3	ATCCACTTTGGTCCCAGGGCCG
JK5	CTCCAGTCGTGTCCCTTGGCCG
GKC	CCTTCCACTGTACTTTGGCCTC

PCR2 primer ID	Primer sequence (first 10 nucl. MID tag)
MID1 VK1	acgagtgcgtCATCCAGWTGACCCAGTCTCC
MID1 VK2	acgagtgcgtGATATTGTGATGACCCAGWCT
MID1 VK3	acgagtgcgtGACRCAGTCTCCAGCCACCCTG
MID1 VK4	acgagtgcgtGACATCGTGATGACCCAGTCT
MID1 VK5	acgagtgcgtGAAACGACACTCACGCAGTCT
MID1 VK6	acgagtgcgtGAAATTGTGCTGACTCAGTCT
MID1 JK1/4	acgagtgcgtTCCACCTTGGTCCCTYSGCCG
MID1 JK2	acgagtgcgtCTCCAGCTTGGTCCCCTGGCCA
MID1 JK3	acgagtgcgtATCCACTTTGGTCCCAGGGCCG
MID1 JK5	acgagtgcgtCTCCAGTCGTGTCCCTTGGCCG
MID1 IGKC	acgagtgcgtCCTTCCACTGTACTTTGGCCTC
MID2 VK1	acgctcgacaCATCCAGWTGACCCAGTCTCC
MID2 VK2	acgctcgacaGATATTGTGATGACCCAGWCT
MID2 VK3	acgctcgacaGACRCAGTCTCCAGCCACCCTG
MID2 VK4	acgctcgacaGACATCGTGATGACCCAGTCT
MID2 VK5	acgctcgacaGAAACGACACTCACGCAGTCT
MID2 VK6	acgctcgacaGAAATTGTGCTGACTCAGTCT
MID2 JK1/4	acgctcgacaTCCACCTTGGTCCCTYSGCCG
MID2 JK2	acgctcgacaCTCCAGCTTGGTCCCCTGGCCA
MID2 JK3	acgctcgacaATCCACTTTGGTCCCAGGGCCG
MID2 JK5	acgctcgacaCTCCAGTCGTGTCCCTTGGCCG
MID2 IGKC	acgctcgacaCCTTCCACTGTACTTTGGCCTC
MID3 VK1	agacgcactcCATCCAGWTGACCCAGTCTCC
MID3 VK2	agacgcactcGATATTGTGATGACCCAGWCT
MID3 VK3	agacgcactcGACRCAGTCTCCAGCCACCCTG
MID3 VK4	agacgcactcGACATCGTGATGACCCAGTCT
MID3 VK5	agacgcactcGAAACGACACTCACGCAGTCT
MID3 VK6	agacgcactcGAAATTGTGCTGACTCAGTCT
MID3 JK1/4	agacgcactcTCCACCTTGGTCCCTYSGCCG
MID3 JK2	agacgcactcCTCCAGCTTGGTCCCCTGGCCA
MID3 JK3	agacgcactcATCCACTTTGGTCCCAGGGCCG
MID3 JK5	agacgcactcCTCCAGTCGTGTCCCTTGGCCG
MID3 IGKC	agacgcactcCCTTCCACTGTACTTTGGCCTC

PCR2 primer ID	Primer sequence (first 10 nucl. MID tag)
MID4 VK1	agcactgtagCATCCAGWTGACCCAGTCTCC
MID4 VK2	agcactgtagGATATTGTGATGACCCAGWCT
MID4 VK3	agcactgtagGACRCAGTCTCCAGCCACCCTG
MID4 VK4	agcactgtagGACATCGTGATGACCCAGTCT
MID4 VK5	agcactgtagGAAACGACACTCACGCAGTCT
MID4 VK6	agcactgtagGAAATTGTGCTGACTCAGTCT
MID4 JK1/4	agcactgtagTCCACCTTGGTCCCTYSGCCG
MID4 JK2	agcactgtagCTCCAGCTTGGTCCCCTGGCCA
MID4 JK3	agcactgtagATCCACTTTGGTCCCAGGGCCG
MID4 JK5	agcactgtagCTCCAGTCGTGTCCCTTGGCCG
MID4 IGKC	agcactgtagCCTTCCACTGTACTTTGGCCTC
MID5 VK1	atcagacacgCATCCAGWTGACCCAGTCTCC
MID5 VK2	atcagacacgGATATTGTGATGACCCAGWCT
MID5 VK3	atcagacacgGACRCAGTCTCCAGCCACCCTG
MID5 VK4	atcagacacgGACATCGTGATGACCCAGTCT
MID5 VK5	atcagacacgGAAACGACACTCACGCAGTCT
MID5 VK6	atcagacacgGAAATTGTGCTGACTCAGTCT
MID5 JK1/4	atcagacacgTCCACCTTGGTCCCTYSGCCG
MID5 JK2	atcagacacgCTCCAGCTTGGTCCCCTGGCCA
MID5 JK3	atcagacacgATCCACTTTGGTCCCAGGGCCG
MID5 JK5	atcagacacgCTCCAGTCGTGTCCCTTGGCCG
MID5 IGKC	atcagacacgCCTTCCACTGTACTTTGGCCTC
MID6 VK1	atatcgcgagCATCCAGWTGACCCAGTCTCC
MID6 VK2	atatcgcgagGATATTGTGATGACCCAGWCT
MID6 VK3	atatcgcgagGACRCAGTCTCCAGCCACCCTG
MID6 VK4	atatcgcgagGACATCGTGATGACCCAGTCT
MID6 VK5	atatcgcgagGAAACGACACTCACGCAGTCT
MID6 VK6	atatcgcgagGAAATTGTGCTGACTCAGTCT
MID6 JK1/4	atatcgcgagTCCACCTTGGTCCCTYSGCCG
MID6 JK2	atatcgcgagCTCCAGCTTGGTCCCCTGGCCA
MID6 JK3	atatcgcgagATCCACTTTGGTCCCAGGGCCG
MID6 JK5	atatcgcgagCTCCAGTCGTGTCCCTTGGCCG
MID6 IGKC	atatcgcgagCCTTCCACTGTACTTTGGCCTC

PCR2 primer ID	Primer sequence (first 10 nucl. MID tag)
MID7 VK1	cgtgtctctaCATCCAGWTGACCCAGTCTCC
MID7 VK2	cgtgtctctaGATATTGTGATGACCCAGWCT
MID7 VK3	cgtgtctctaGACRCAGTCTCCAGCCACCCTG
MID7 VK4	cgtgtctctaGACATCGTGATGACCCAGTCT
MID7 VK5	cgtgtctctaGAAACGACACTCACGCAGTCT
MID7 VK6	cgtgtctctaGAAATTGTGCTGACTCAGTCT
MID7 JK1/4	cgtgtctctaTCCACCTTGGTCCCTYSGCCG
MID7 JK2	cgtgtctctaCTCCAGCTTGGTCCCCTGGCCA
MID7 JK3	cgtgtctctaATCCACTTTGGTCCCAGGGCCG
MID7 JK5	cgtgtctctaCTCCAGTCGTGTCCCTTGGCCG
MID7 IGKC	cgtgtctctaCCTTCCACTGTACTTTGGCCTC
MID8 VK1	ctcgcgtgtcCATCCAGWTGACCCAGTCTCC
MID8 VK2	ctcgcgtgtcGATATTGTGATGACCCAGWCT
MID8 VK3	ctcgcgtgtcGACRCAGTCTCCAGCCACCCTG
MID8 VK4	ctcgcgtgtcGACATCGTGATGACCCAGTCT
MID8 VK5	ctcgcgtgtcGAAACGACACTCACGCAGTCT
MID8 VK6	ctcgcgtgtcGAAATTGTGCTGACTCAGTCT
MID8 JK1/4	ctcgcgtgtcTCCACCTTGGTCCCTYSGCCG
MID8 JK2	ctcgcgtgtcCTCCAGCTTGGTCCCCTGGCCA
MID8 JK3	ctcgcgtgtcATCCACTTTGGTCCCAGGGCCG
MID8 JK5	ctcgcgtgtcCTCCAGTCGTGTCCCTTGGCCG
MID8 IGKC	ctcgcgtgtcCCTTCCACTGTACTTTGGCCTC
MID9 VK1	tagtatcagcCATCCAGWTGACCCAGTCTCC
MID9 VK2	tagtatcagcGATATTGTGATGACCCAGWCT
MID9 VK3	tagtatcagcGACRCAGTCTCCAGCCACCCTG
MID9 VK4	tagtatcagcGACATCGTGATGACCCAGTCT
MID9 VK5	tagtatcagcGAAACGACACTCACGCAGTCT
MID9 VK6	tagtatcagcGAAATTGTGCTGACTCAGTCT
MID9 JK1/4	tagtatcagcTCCACCTTGGTCCCTYSGCCG
MID9 JK2	tagtatcagcCTCCAGCTTGGTCCCCTGGCCA
MID9 JK3	tagtatcagcATCCACTTTGGTCCCAGGGCCG
MID9 JK5	tagtatcagcCTCCAGTCGTGTCCCTTGGCCG
MID9 IGKC	tagtatcagcCCTTCCACTGTACTTTGGCCTC

Table 2.3 Sequences of MID tags and PCR primers used for second round amplification of *IGK* rearrangements for the HTS analysis.

2.3.3 cDNA synthesis and PCR amplification of gDNA and cDNA

Isolation of genomic DNA from populations of sorted cells is described in Section 2.5 of this materials and methods chapter. Vκ-Jκ amplifications were performed from all gDNA samples.

To synthesise cDNA from sorted cell populations, cells were centrifuged at 1200rpm for 5 minutes and the supernatant removed. Cells were re-suspended in 180µl of SLyRT buffer in a PCR tubes and gently mixed. 20µl of 1:8 500U SuperScript III reverse transcriptase (RT; Invitrogen) was then added to the cell/SLyRT suspension. The reverse transcription was performed at 42°C (10 minutes), 25°C (10 minutes), 50 °C (60 °C minutes) and 72°C (15 minutes). Vκ-Jκ and Vκ-Cκ rearrangements were amplified using semi-nested polymerase chain reaction (PCR). In PCR1, a 25µL reaction mix contained 6.25µL of cDNA, 0.625U Phusion DNA polymerase (NEB), 200µM each dNTPs, 41.75nM each upstream *IGKV* primers, 250nM downstream *IGKC* primer in 1× reaction buffer. After a hot start at 98 °C (for 30 seconds, hold at 50 °C) Phusion DNA polymerase was added, followed by 15 cycles of 98 °C, (10 seconds); 58 °C (15 seconds); 72 °C (30 seconds), and 1 cycle of 72 °C (5 minutes).

PCR2 was then used to amplify from PCR1 products using primers having 10-base multiplex-identifier (MID) tails. MID tags enabled 9 different samples to be pooled into 1 sequencing sample and individual experimental samples were later separated by sequence analysis of the MID tags. Twenty µl of PCR2 reaction mix contained 2 µl of PCR1 products, 0.5U Phusion DNA polymerase, 200µM each dNTPs, 41.75nM each upstream MID: *IGKV1-6* primers, 250nM *IGKC* downstream primer in 1× reaction buffer. PCR2

was performed at: 98 °C for (30 seconds), 20 cycles of 98 °C (10 seconds); 58 °C (15 seconds); 72 °C (30 seconds), and 1 cycle of 72 °C (5 minutes). Recipe summaries for each PCR reaction are provided in Table 2.4 and Table 2.5.

Table 2.4 Recipes and concentrations for HTS first round PCR reactions (without MID tags)

Reagents	**Stock conc.**	**µl/reaction**	**Final conc.**
GC buffer	5X	5	1X
1:40 Phusion DNA polymerase	0.125 U/µl	5	0.625 U
dNTPs	20 mM each	0.25	200 µM each
Vκ1-6 (5') primer mix	835 nM each	1.25	41.75 nM each
VκC (3') primer mix	5 µM each		+ 250 nM
PCR H_20	-	7.25	-
cDNA	-	6.25	~200 ng
Volume	-	**25**	

Table 2.5 Recipes and concentrations for HTS second round PCR reactions (with MID tags.)

Reagents	**Stock conc.**	**µl/reaction**	**Final conc.**
GC buffer	5X	4	1X
1:20 Phusion DNA polymerase	0.25 U/µl	2	0.5 U
dNTPs	20 mM each	0.2	200 µM each
MID: Vκ1-6 (5') primer mix	835 nM each +	1	41.75 nM each
MID: VκC (3') primer mix	5 µM each		+ 250 nM
PCR H20	-	10.8	-
cDNA	-	2	-
Volume	-	**20**	

2.3.3 Preparation of MID-tagged PCR products

2.3.3.1 General information

In order to obtain sufficient quantities of PCR product for sequencing, first round PCR reactions were performed eight times for each DNA sample, from which second round PCR reactions were carried out in duplicate (final volume per DNA sample: 20 µl x 16 = 320 µl.) Total PCR products from 16 second round PCRs of the same MID tag were pooled and heated to 95°C for 5 minutes.

2.3.3.2 Gel electrophoresis of MID-tagged PCR products and purification

For the running of PCR products for high-throughput sequencing, 1x TAE buffer was used to prepare a 1.5% agarose gel consisting of 4.5g 3:1 agarose (Flowgen) dissolved in 300ml 1x TAE (Recipe section 2.2.2.1). PCR products were run for 40 minutes at 120V and bands of target size were extracted using QIAquick Gel Purification Kit (Qiagen).

After determining the concentration of each pooled DNA sample, equal quantities of purified PCR products (in ng) from 9 different MID types for each experimental set (control and SLE) were passed through QIAquick spin columns to remove any primer carryover and to concentrate samples to 50 ng/µl. A minimum of 5 µg purified DNA was required. The sequencing was carried out on the GS FLX System by Dr. Gerald Hyakatura and Dr. Annett Helfrich-Kröger at LGC Genomics GmBH.

2.3.4 Sequence analysis

Sequences were assigned to the corresponding samples based on the terminal MID sequence. Sequences that contained a second MID sequence that was either different, or

located internally, were excluded. A series of stringent quality control criteria were applied to exclude biologically implausible sequences. *IGKV-C* rearrangements amplified from cDNA were only accepted as biologically plausible if they were over 415 nucleotides in length and Vκ-Jκ rearrangements amplified from the DNA of Igκ- or Igλ-expressing B cells were accepted providing they were over 290 nucleotides in length. Sequences that failed to start with a MID tag, *IGKV* with either *IGKJ* or *IGKC* were rejected, as were sequences with multiple *IGKV* gene primer motifs. Identical sequences were identified by sorting based on their CDR3 nucleic acid sequences with *IGKV* and *IGKJ* gene use as secondary identifiers and only one included in the data.

Comparisons involved analysis of contributions of individual *IGKV* and *IGKJ* segments to the total pool of rearrangements. Since rearrangements of 30 different *IGK* segments were observed in a single data set, groups of sequences were required to contain 300 different rearrangements for a single comparison of *IGKV* segment usage. Data from different individual donors were therefore pooled to give a larger dataset for accuracy. Where it was useful for verification using datasets from individual donors we determined that a single donor could be analysed alone provided the set for analysis contained more than 300 different rearrangements.

Analyses were performed in Excel (Microsoft). Proportions were compared using χ^2 tests with Bonferroni correction for multiple comparisons.

2.4 Flow cytometry

2.4.1 Protocol

PBMCs were thawed following liquid N_2 storage in 1 ml FCS. Following a 5 minute wash in 15ml RPMI-1640 medium cells were counted and immunostained in 100µl 2% FCS buffer and the appropriate volumes of antibody (See Table 2.6). Cells were stained on ice for 20 minutes in the dark.

Table 2.6 Mouse anti-human antibodies used for flow cytometry analysis and sorting

Mouse anti-Human Abs	Fluorochrome	vol. in 100 µl 2%FCS buffer	Cell no. /stain vol	Isotypes	Manufact.
CD19	PerCp-Cy5.5	5µl/mil cells	< 10million	Mouse IgG1 κ PerCP-Cy5.5	BD Pharm
IgD	PE	6µl/mil cells		Mouse IgG1 κ PE	BD Pharm
CD27	PE-Cy7	5µl/mil cells		Mouse IgG1 κ PE-Cy7	BD Pharm
CD10	APC	5µl/mil cells		Mouse IgG1 k APC	BioLegend
CD38	PE	5µl/mil cells		Mouse IgG1 k PE	eBioscience
IgA	FITC	10µl/mil cells		Mouse IgG1 κ FITC	Miltenyi Biotech
Igκ	APC-H7	5µl /mil cells		Mouse IgG1 κ APC-H7	BD Pharm
Igλ	Pacific Blue	3µl /mil cells		Mouse IgG2a κ APC-H7	BioLegend
VpreB (CD179a)	PE	5µl /mil cells		Mouse IgG1 κ PE	BioLegend
Viability stain	Live/Dead Fixable Aqua	1µl /mil cells	N.A	N.A	Life Technologies /Invitrogen

Fluorescence minus one (FMO) controls and appropriate single and isotype controls were also stained alongside the cell samples. Titration experiments had previously been performed on antibodies to verify the optimal staining concentrations. Following cell surface staining, cells were washed in 2ml 2% FCS buffer for 5 mins at 1200 rpm and 4°C. A wash step at the same spin conditions but using 2ml 1x PBS was performed prior to viability staining with the Live/Dead cell viability stain. All cells were stained for viability on ice in 100ml 1x PBS for 30 minutes in the dark. Following staining, cells were washed

in 2ml 2% FCS and cell pellets were re-suspended in 200ml 2%FCS. Cells for sorting were sorted into polypropylene FACS tubes containing 1.5ml of 2% FCS buffer. Upon completion of cell sorting, cells were centrifuged for 5 minutes at 300g and the supernatant removed.

2.4.2 Analysis

For flow cytometry analysis without sorting, cells were acquired and recorded using the FACSCanto (BD). Cell sorting was performed using the FACSAria-II (BD) with sort purities >90%. All flow cytometry experiments could be subsequently analysed on Flow Jo (V7.5).

2.4.2.1 Subset phenotypic characterization

The subsets of B cells studied in this thesis are represented in Table 2.7. In most instances the gating of these B cell populations was followed by separation of light chain expression, kappa and lambda.

Table 2.7 Human B cell subset phenotypic characterisation

B cell subset	Cell surface antigen expression levels
Total B cells	CD19+
Transitional B cells	CD19+ IgD+ CD27- CD10hi
Mature naive B cells	CD19+ IgD+ CD27- CD10-
Memory B cells	CD19+ IgD- CD27+/- CD10-
IgA plasmablasts	CD19+ CD38+ CD27hi sIgA+
Non-IgA plasmablasts	CD19+ CD38+ CD27hi sIgA-

2.5 Preparation of DNA, RNA and cDNA samples from populations of sorted cells

DNA isolated from FACS-sorted populations of B cells was performed using the DNeasy Blood and Tissue Kit (Qiagen) where DNA was eluted in 200µl elution buffer before storage at -20°C prior to use. Isolated DNA from sorted B cell populations was used for all qRT-PCR analyses in this study, presented in Chapter 4, and was similarly performed prior to sequence analysis of iRSS-KDE junctional analyses (Chapter 4) and VκJκ rearrangements (Chapter 5). RNA samples were obtained through isolated of RNA from FACS-sorted $CD19^+$ B cells using the miRNeasy Mini Kit (Qiagen). To prevent contamination of RNA preps during isolation, all equipment to be used was exposed to UV light for 20 minutes in a hood designated to RNA isolation. Newly extracted RNA samples were immediately stored at -80°C and were kept on ice during preparation of reverse transcription.

2.5.1 Reverse transcription of RNA isolated from sorted B cell populations

Reverse transcription of isolated RNA samples to cDNA was performed using Oligo dT primer and an M-MLV Reverse Transcriptase kit (both Promega). Each reverse transcription contained the following volumes of RT reagents: 7µl RNA sample, 3µl molecular standard PCR H_20 and 1µl Oligo dT primer. Reagents were mixed gently prior to incubation at 70°C for 10 minutes. RT reactions were subsequently removed and allowed to cool on ice for 10 minutes before centrifuging for 1 minute to collect any products that had condensed around the PCR tubes during incubation. Following this, the addition of a 4µl M-MLV buffer, 0.5µl dNTPs, 0.5µl RNase inhibitor, 1µl M-MLV Reverse

Transcriptase (all Promega) and 3 µl dH_20 were added to give a final volume of 20µl. The reactions were incubated for a further 1 hour at 42 °C. A final extension period of 15 minutes at 70°C was applied to the samples and held at 15 °C before use.

2.6 Real time-PCR

2.6.1 General information

Quantitative real-time PCR for the amplification of rearranged iRSS-KDE junctions were designed by Applied Biosystems using a Taqman Gene Expression Assay. iRSS-KDE rearrangements were amplified from gDNA with 5'-ATTGATGCTGCCGTAGCC-3' and 5'-AGGCTTCCTAGGGAGGTCAG-3' primers and were detected with the FAM-labelled 5'-TCTGCAGCTGCATTTTTGCCA-3' hydrolysis probe (Panigrahi et al., 2008). All qRT-PCRs were performed in multiplex, with quantification of iRSS-KDE rearrangements standardised to the Taqman Copy Number Reference Assay, RNaseP, detected with a VIC-labelled probe (Applied Biosystems). Quantitative real-time PCR for the amplification of an intronic region 5' to the Jκ-Cκ intronic RSS were also produced by Applied Biosystems using a similar Taqman Gene Expression Assay. This region was amplified with the following primers 5'-GGGTCTGATGGCCAGTATTGAC-3' and 5'-CCAATACAATATTTGGCAAG-3' and were detected with a FAM-labelled 5'-AAAGATTGGTAAATGAGGGCATTTA-3' hydrolysis probe. This assay was also performed in multiplex real-time PCR reaction with the RNaseP reference assay.

2.6.2 Protocol

Quantitative real-time PCRs were carried out in 384-well plates in a 10µl reaction volume. 4µl of sample gDNA was added to 6µl reaction master mix containing 5µl 2x Gene Expression Master Mix (Applied Biosystems); 0.5µl iRSS-KDE primer and 0.5µl RNaseP reference primer. KDE-RSS rearrangements were amplified under the following quantitative PCR conditions: denaturation at 95°C for 10 mins, followed by 40 cycles of: 95°C for 15 secs followed by 60°C for 60 secs. Individual samples were performed in triplicate, and the reliability of replicates and calculations performed to quantify iRSS-KDE rearrangements will be described in Chapter 4.

2.7 Amplification of Ig gene rearrangements

2.7.1 General information

This study involved the analysis of VκJκ rearrangements and iRSS-KDE rearrangements from the genomic DNA of $CD19^+$ $Ig\kappa^+/Ig\lambda^+$ FACS-sorted B cells. Both VκJκ and iRSS-KDE rearrangements were amplified by nested PCR reactions. Each PCR reaction contained the following quantities of reagents: 10µl 5x PCR buffer, 3µl $MgCl_2$, 0.5µl each 5'/3' primer (20mM), 0.5µl 20mM dNTPs, 3µl DNA. PCR reactions were amplified in a total volume of 50ml, so dH_20 was added to the PCR mix depending on the total number of 5' and 3' primers used.

2.7.2 PCR cycle conditions

- VκJκ and KDE-JκCκ intronic RSS rearrangements :
    - Round 1: One cycle of 95°C for 7 min, 56°C for 1 min and 72°C for 1 min 30s. Followed by 30 cycles of 94°C for 1 min, 56°C for 30 secs and 72°C for 1 min 30s. A final extension of 5 mins at 72°C was performed.
    - Round 2 (Nested PCR): 5 ml of each first round PCR product was used. One cycle of 95°C for 7 min, 65°C for 1 min and 72°C for 1 min 30s. Followed by 30 cycles of 94°C for 1 min, 65°C for 30 secs and 72°C for 1 min 30s. A final extension of 5 mins at 72°C was performed.

2.7.3 PCR primers

The following table provides the primer sequences used for the amplification of VκJκ rearrangements and iRSS-KDE rearrangements.

Table 2.8 Sequences of oligonucleotide primers for the amplification of VκJκ rearrangements (Foster et al., 1997)

Primer name	**Sequence 5' to 3'**
External 5' primers	
VκI/IIE	GCTCAGCTCCTGGGGCT
VκIIIE	GGAA(AG)CCCCAGC(AGT)CAGC
VκIV/VE	CT(CG)TT(GC)CT(CT)TGGATCTCTG
VκVI/VIIE	CT(GC)CTGCTCTGGG(CT)TCC
External 3' primers	
Jκ2E	ACGTTTGATCTCCAGCTTG
Jκ5E	CTTACGTTTAATCTCCAGTC
Internal 5' primers	
VκI-I	CATCCAG(AT)TGACCCAGTCTCC
VκII-I	TCCAGTGGGGATATTGTGATGAC
VκIII-I	GTCT(GT)TGTCTCCAGGGGAAAGAG
VκIV-I	GACATCGTGATGACCCAGTCTC
VκV-I	GGGCAGAAACGACACTCACGCA
VκVI-I	TCCAGGGGTGAAATTGTG(AC)TGAC
VκVII-I	GCTGCAATGGGACATTGTGCT
Internal 3' primers	
Jκ2I	CAGCTTGGTCCCCTGGCCAAA
Jκ5I	CCAGTCGTGTCCCTTGGCCG

Table 2.9 Sequences of oligonucleotide primers for the amplification of KDE rearrangements to the Jκ-Cκ intronic RSS.

Primer name	Sequence 5' to 3'
External 5' primer	
JκCκ intron E	CCGCGCTCTTGGGGCAGCCGCC
External 3' primer	
KDE E	CAGACAGGTCCTCAGAGGTC
Internal 5' primer	
JκCκ intron I	GCCGCTATGCCGTGGCCACCC
Internal 3' primer	
KDE I	CAGGTTGTCCTAGGGAGCAGGG

Table 2.10 Sequences of oligonucleotide primers for the amplification of surrogate light chain VpreB.

Primer name	Sequence 5' to 3'
External 5' primer	
VpreB E	CATGCTGTTTGTCTACTGCACAG
External 3' primer	
VpreB 5	TGCAGTGGGTTCCATTTCTTCC
Internal 5' primer	
VpreB I	GCCGCCGGCCATGTCCTCGGCCC
Internal 3' primer	
VpreB I	CCTCCTTCTCCGAGCTGCGGGC

2.7.3 Separation of PCR products by gel electrophoresis

PCR products from amplifications not involved in the high-throughput sequencing were run on a 1.5% agarose gel prepared with 3g of MicroSieve 3:1 agarose (Flowgen) dissolved

in 200ml 1x TBE solution by a 4 minute microwave heat on a medium setting. Following cooling and prior to pouring into the gel mould, 1 µl 50mg/ml ethidium bromide solution was swirled into the gel, which was allowed to set for 30 minutes. 6 µl of each PCR product was then mixed with 2 µl of gel loading dye and loaded into a single well. 5 µl of ΦX174 DNA11Hinf1 marker (Promega) was run alongside PCR products for identification of the PCR product size following the gel run at 125 V for 45 minutes in 1x TBE. Finally gels were visualized under UV light.

2.7.4 Purification of PCR products

PCR products were purified with the MinElute PCR purification kit (Qiagen) and protocol followed to manufacturer's instructions. Samples were stored at -20°C until use.

2.8 Molecular cloning of PCR products

2.8.1 Ligation

The pGEM-T vector kit (Promega) was used to ligate PCR products on ice. The ligation reaction was performed in 10µl; 3µl purified PCR products, 5µl ligation buffer (10X), 1µl of T4 DNA ligase (3µ/ml) and 1µl of vector and incubated at RT for 2 hours, or at 4°C overnight.

2.8.2 Transformation

2.5 µl of each ligation reaction and 25µl competent cells (JM109, Promega) were added to a 1.5ml sterilised Eppendorf and incubated on ice for 20 minutes. The bacteria were heat shocked in a water bath at 42°C for 45 seconds and incubated immediately on ice for 2

minutes. 950 µl of cold SOC medium was added to each tube and incubated on a shaker (150 rpm) at 37°C for 2 hours. Bacteria were collected by centrifuging at 200 rpm at RT for 10 minutes and the supernatant discarded. The bacterial pellet was re-suspended in 100µl of cold SOC medium, and 100µl was spread on individual plates prepared with LB agar containing 0.5% ampicillin. 5% X-Gal (Promega) and 0.05M IPTG. Plates were inverted and incubated at 37°C overnight.

2.8.3 Blue/white screening of clones

Bacteria successfully transformed with pGEM-T vector containing the rearranged KDE insert grew as white colonies on the plates. Bacteria that failed to be transformed grew as blue colonies. Discrete white colonies were picked randomly with sterile pipettes, transferred to a griddled plate and allowed to grow overnight at 37°C. Pipette tips used to select clones were rinsed into 15µl dIH_20 in individual wells of a 96-well plate. The plates were covered and placed in a PCR machine to boil for 10 minutes. This releases vectors and denatures the DNases present. The plate was centrifuged for 10 minutes at maximum speed to pellet the vector and stored at -20°C.

2.8.4 Screening clones by PCR

5µl of the supernatant of screened clones was used as template for a PCR reaction containing 1.5mM $MgCl_2$, 0.2M dNTPs, 1x Taq DNA polymerase reaction buffer, 50ng of each M13For and M13Rev primers and 0.25 units Taq DNA polymerase in a total volume of 10µl. After a hot start for 5 minutes at 95°C, 30 cycles of 95°C for 30 seconds, 45°C for 30 seconds and 72°C for 90 seconds, were followed by a 5 minute extension phase at 72°C.

PCR products were visualised after running on a 1.5% agarose gel (previously described) and products of the correct size were identified.

2.8.4.1 PCR cycle conditions for clone screen

Samples were denatured for 5 minutes at 95°C, followed by 30 cycles of: 95°C for 30 secs, 45°C for 30 secs and 72°C for 1 min 30 s. A final extension of 5 minutes at 72°C was performed.

Table 2.11 Sequences of M13 oligonucleotide primers used for the screening of transformed clones

Primer name	Sequence 5' to 3'
M13 Fwd	GTAAAACGACGGCCAGT
M13 Rev	GGAAACAGCTATGACCATG

2.8.5 DNA sequencing

Clones were selected from the griddled plate according to the size of the cloned PCR products and transferred to individual wells of a 96-well plate containing LB agar + 0.1m/ml ampicillin. Plates were incubated at 37°C overnight before sending to Beckman and Coulter for sequencing.

2.9 RNA isolation and reverse transcription from single cells

Cells from dual-light chain-expressing SLE patient Lu-025 were FACS-sorted into 96-well plates containing 18µl of SLyRT buffer per well. Plates were then immediately covered with a plastic sterile lid and spun briefly in the centrifuge and kept on ice. In a clean PCR hood, the lid was gently removed and 2µl of 1 in 8 SuperScript III (Invitrogen) was added

to each PCR well before a further brief centrifuge spin to ensure each transcription reaction was contained within the bottom of each well. Plates were then reverse transcribed under the following conditions; denaturation at 42°C for 10 mins, annealing at 25°C for 10 mins, an extension period at 55°C for 60 mins followed by termination of transcription at 72°C for 15 mins. Plates could then be stored at -20°C until further use.

2.10 PCR amplification of rearranged *IGK* (VκCκ) and *IGL* (VλCλ) transcripts from cDNA isolated from single cells.

2.10.1 General information

Semi-nested PCRs were used to detect and amplify *V-C* transcripts for both *IGK* and *IGL* light chain gene rearrangements in the cDNA of sorted single cells.

Table 2.12 Recipes and concentrations for first round semi-nested PCR amplification of *IGK* and *IGL* transcripts from single cells

Reagents	**Stock conc.**	**µl/reaction**
GC buffer	5X	4
1:20 Phusion DNA polymerase	0.25 U/µl	2
dNTPs	20 mM each	0.2
Vκ1-6 or Vλ1-8(5') primer mix for PCR 1	835 nM or 1000nM each	1
VκC or VλC (3') primer mix for PCR 1	5 µM each	
PCR H20	-	8.8
MgCl2	2.0 mM	1
cDNA	-	3
Volume	-	**20**

Table 2.13 Recipes and concentrations for first second round semi-nested PCR amplification of VκCκ and VλCλ transcripts from single cells

Reagents	**Stock conc.**	**µl/reaction**
GC buffer	5X	4
1:20 Phusion DNA polymerase	0.25 U/µl	2
dNTPs	20 mM each	0.2
Vκ1-6 or Vλ1-6(5') primer mix for PCR 2	835 nM or 1000nM each	1
VκC or VλC (3') primer mix for PCR 2	5 µM each	
PCR H20	-	10.8
PCR1 product	-	2
Volume	-	**20**

2.10.2 PCR cycle conditions

- **Vκ-Cκ:**
  - Round 1: 98 for 30secs, followed by 50 cycles of 98°C for 10 secs, 58°C for 15 secs and 72°C for 40 secs. Final extension at 72°C for 10 mins.
  - Round 2: 98°C for 30secs, followed by 20 cycles of 98°C for 10 secs, 64°C for 15 secs and 72°C for 15 secs. Final extension at 72°C for 10 mins.
- **Vλ-Cλ:**
  - Round 1: 98 for 30secs, followed by 50 cycles of 98°C for 10 secs, 58°C for 15 secs and 72°C for 30 secs. Final extension at 72°C for 10 mins.
  - Round 2: 98°C for 30secs, followed by 20 cycles of 98°C for 10 secs, 62°C for 15 secs and 72°C for 15 secs. Final extension at 72°C for 10 mins.

2.10.3 PCR primers for semi-nested PCR amplification of VκCκ and VλCλ from single cells

Table 2.14 Sequences of oligonucleotide primers for the first round amplification of VκCκ and VλCλ transcripts from single cells

PCR 1	Primer name	Primer sequence 5' to 3'
Vκ-Cκ ***external***	VK1	CATCCAGWTGACCCAGTCTCC
	VK2	GATATTGTGATGACCCAGWCT
	VK3	GACRCAGTCTCCAGCCACCCTG
	VK4	GACATCGTGATGACCCAGTCT
	VK5	GAAACGACACTCACGCAGTCT
	VK6	GAAATTGTGCTGACTCAGTCT
	CK1N	TTAACACTCTCCCCTGTTGAAGC
Vλ-Cλ ***external***	LVL1	GGTCCTGGGCCCAGTCTGTGCTG
	LVL2	GGTCCTGGGCCCAGTCTGCCCTG
	LVL3	GCTCTGTGACCTCCTATGAGCTG
	LVL4/5	GGTCTCTCTCSCAGCYTGTGCTG
	LVL6	GTTCTTGGGCCAATTTTATGGTG
	LVL7	GTTCTTGGGCCAATTTTATGCTG
	LVL8	GAGTGGATTCTCAGACTGTGGTG
	IGLC outer	CCTTCCAGGCCACTGTCAC

Table 2.15 Sequences of oligonucleotide primers for the second round amplification of VκCκ and VλCλ transcripts from single cells

PCR2	Primer name	Primer sequence 5' to 3'
Vκ-Cκ ***internal***	VK1	CATCCAGWTGACCCAGTCTCC
	VK2	GATATTGTGATGACCCAGWCT
	VK3	GACRCAGTCTCCAGCCACCCTG
	VK4	GACATCGTGATGACCCAGTCT
	VK5	GAAACGACACTCACGCAGTCT
	VK6	GAAATTGTGCTGACTCAGTCT
	IGKC2	CTGTACTTTGGCCTCTCTGGG
Vλ-Cλ ***internal***	LVL1	GGTCCTGGGCCCAGTCTGTGCTG
	LVL2	GGTCCTGGGCCCAGTCTGCCCTG
	LVL3	GCTCTGTGACCTCCTATGAGCTG
	LVL4/5	GGTCTCTCTCSCAGCYTGTGCTG
	LVL6	GTTCTTGGGCCAATTTTATGGTG
	LVL7	GTTCTTGGGCCAATTTTATGCTG
	LVL8	GAGTGGATTCTCAGACTGTGGTG
	IGLC2	GCCACTGTCACRGCTCCCGGGC

Chapter 3
High throughput sequencing of the human mature naive immunoglobulin kappa light chain repertoire

3.1 Introduction

This chapter describes an analysis of the *IGK* light chain repertoire of mature naive B cells in health and SLE. The introduction to this chapter will be divided into background, experimental design and experimental aims.

3.1.1 Background

Systemic lupus erythematosus (SLE) is an autoimmune disease of unknown origin, characterised by chronic B cell hyper-activation and an autoimmune immunoglobulin repertoire that targets DNA and nucleoproteins (Dorner et al., 2011a). The breakage of tolerance to self-antigen in SLE is thought to be a consequence of failure of an unknown systemic checkpoint that governs the transition of B cells from the early bone marrow emigrant to a mature naive stage in development (Yurasov et al., 2005). A relatively large proportion of functional immunoglobulin expressed by early emigrant B cells expresses varying degrees of polyreactivity and autoreactivity (Mouquet and Nussenzweig, 2012; Wardemann et al., 2003). In health, a series of checkpoints serve to counter-select B cells with self-reactive immunoglobulins during B cell development. At these checkpoints, B cells can be removed from the developing repertoire by several mechanisms; clonal deletion, the induction of anergy and receptor editing, which involves secondary VJ rearrangement events on the same allele at the light chain loci (Goodnow et al., 1988; Halverson et al., 2004; Nemazee and Burki, 1989; Radic et al., 1993; Tiegs et al., 1993).

In SLE however, studies investigating a role of receptor editing in disease pathogenesis have provoked debate. It has been reported that more apparent receptor editing is a feature of an individual patient with SLE compared to health by analysis of genomic DNA from

single $CD19^+$ B cells (Dorner et al., 1998). Contrastingly, it has been suggested that receptor editing is reduced in SLE compared to healthy individuals through study of the rearranging activity of the kappa deleting element (Panigrahi et al., 2008).

Diversity in the immunoglobulin repertoire is generated by a hierarchical series of gene rearrangement events initiated during B cell development in the bone marrow (Meffre et al., 2000a). After rearrangement of a functional *IGH* chain, rearrangement on one allele at the *IGK* loci follows (Alt et al., 1984). If the first rearrangement of the first *IGK* allele is successful, the *IGK* light chain is paired with a heavy chain and subsequently expressed on the cell surface. When *IGK* rearrangements are unsuccessful, replacement rearrangement events on the same allele can follow in attempt to produce functional light chains of *IGK*. If unsuccessful, the *IGL* light chain locus is activated until a functional light chain is produced (Panigrahi et al., 2008). The successful pairing of functional heavy and light chains on the cell surface initiates the cellular and molecular processes to regulate central tolerance in the B cell system. The extreme repertoire diversity is the result of several processes in the developing B cell receptor. These involve combinatorial diversity introduced through the rearrangement of different V, (D) and J gene segments, junctional diversity introduced by TdT and exonuclease activity at the segmental junctions, and combinatorial diversity associated with different heavy and light chain pairs (Kawasaki et al., 2001; Kawasaki et al., 1997; Weigert et al., 1980).

As a result of this activity at the light chain loci, B cells isolated from peripheral blood will typically contain several Ig light chain gene rearrangements. These may include rearrangement events at one or both *IGK* alleles and in some B cells, one or two rearrangement events at the *IGL* loci additionally (Bräuninger et al., 2001). Many of these rearrangements will be in the incorrect genetic reading frame and therefore not functional and not subsequently produced as light chain protein. Other rearrangements will be

productive and expressed by the cell. Unwanted light chain rearrangements are typically inactivated by the rearranging activity of the KDE to an RSS located in the Jκ-Cκ intron (Klobeck and Zachau, 1986). This rearrangement event removes the constant region and enhancers flanking it from the rearranged IGK allele rendering it functionally silenced (Siminovitch et al., 1985). In this way, the KDE acts alongside other mechanisms overseeing central tolerance by preventing the expression of more than one light chain, thus maintaining isotypic exclusion. Usage of more than one rearranged *IGK* light chain, or *IGK* and *IGL* light chains by B cells has been shown to have a profound effect on maintenance and breakage of tolerance in animal models of autoimmunity (Andrews et al., 2013; Casellas et al., 2007; Fournier et al., 2012). Skewed usage of IGKV gene segments in SLE has previously been reported.

3.1.2 Experimental design

High-throughput sequencing technologies were originally used in 2005 and have now been developed to allow rapid and large-scale parallel sequencing; a platform now widely used in the field of immunoglobulin repertoire analyses. The Roche FLX Genome Titanium Sequencer (Agowa GmBH) system can sequence hundreds of samples in parallel through the use of sample-specific multiple identifier (MID) tags. This tool has been applied to study the human *IGH* repertoire of transitional, naive and IgM memory and switched memory B cell populations (Wu et al., 2010) and was applied in a study aiming to analyse paired human heavy and light chain immunoglobulin repertoires by isolation of mRNA from single B cells and performing linkage PCR prior to high-throughput sequencing (Dekosky et al., 2013).

3.1.2.1 FACS sorting procedure

In order to investigate the factors that shape the expressed *IGK* light chain repertoire in health and in the context of SLE, a high-throughput sequencing approach was applied to FACS-sorted mature naive B cells ($CD19^+$ IgD^+ $CD27^-$ $Ig\kappa^+/Ig\lambda^+$) from PBMCs isolated from the blood of three SLE patients and three healthy controls. Of note, it became apparent during the FACS-sorting of some SLE samples that populations of single $Ig\kappa^+$ and $Ig\lambda^+$ $CD19^+$ B cells were not as clearly defined as the two discrete populations of single light chain producers typically observed during analysis of healthy samples stained by the same method. To ensure clarity and avoid the interference of an unknown experimental variable, SLE samples were selected for this sequencing approach based on the existence of two discrete populations of cells expressing either Igκ or Igλ on their cell surface.

Genomic DNA was isolated from sorted populations of both Igκ- and Igλ-expressing B cells, and RNA was isolated and reverse transcribed from Igκ-producers only. Two primer combinations were used to amplify rearranged IGK genes in a semi-nested PCR consisting of two rounds. *IGKV* and *IGKJ* primers were used to amplify VJ junctions from the genomic DNA of Igκ- and Igλ-producing B cells, whereas an *IGKC* primer was used in combination with the same *IGKV* 5'-end primers to amplify cDNA isolated from Igκ-producing B cells (Figure 3.1).

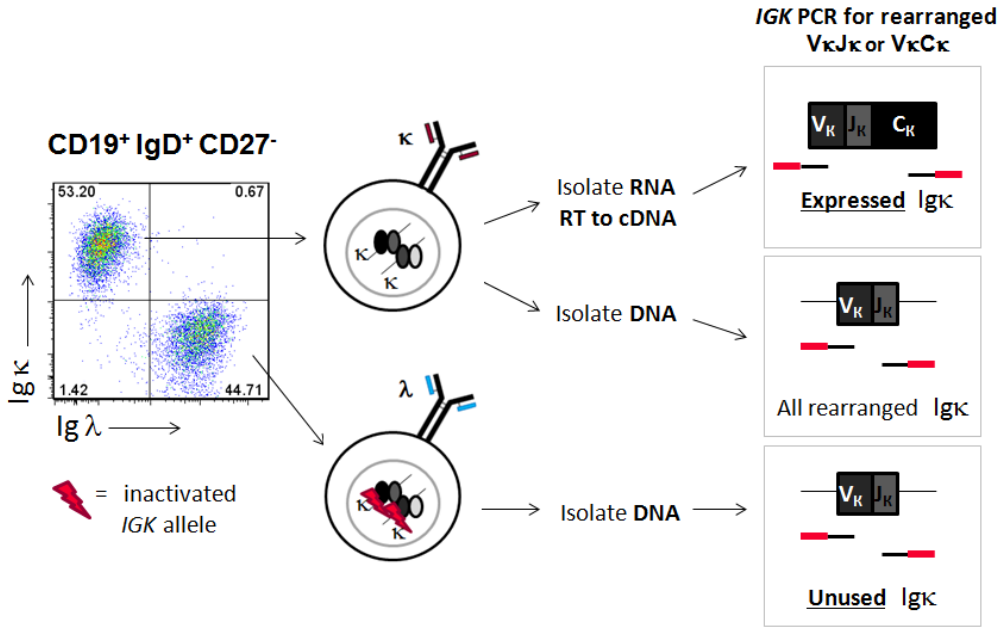


Figure 3.1 Investigating the IGK light chain repertoire in health and SLE. Outline of the experimental setup prior to high-throughput sequencing analysis of *IGK* rearrangements. $CD19^+$ IgD^+ $CD27^-$ Igκ/Igλ-expressing B cell populations were FACS-sorted from total PBMCs of three healthy donors and three SLE patients and genomic DNA isolated from both Igκ- and Igλ-expressing cells for V-J PCR amplification, and RNA was isolated from just Igκ-producers for V-C amplification following reverse transcription to cDNA. The red primer ends represent the MID tags used in the second nested PCR rounds for sample identification.

3.1.2.2 Preparation of PCR-amplified samples

To produce sufficient quantities of DNA for sequencing while maintaining a small number of amplification cycles, multiple sample aliquots were amplified by PCR1 (x8 sample aliquots) and PCR2 (x2 sample aliquots from each PCR1 product) for each experimental sample. PCR2 amplification of PCR1 products used primers with 10-base multiplex-identifier (MID) tails which enabled for 9 different sample conditions to be pooled into one sequencing sample. Healthy controls and SLE samples were kept separate throughout the experiment so MID tags 1-9 could be used for both the 3 controls and 3 SLE patient samples

for analysis. PCR primers were removed from pooled PCR products from each experimental sample (16 final PCR-amplified aliquots) by gel electrophoresis and using a QIAquick Gel Purification Kit. Samples to be sent for sequencing were quantified and mixed in equal quantities before concentrating using QIAquick PCR Purification Kit and being sequenced on the GS FLX Genome Titanium Sequencer (Agowa GmBH.) Individual experimental samples could be recognised after sequencing through identification of their terminal MID sequence (Figure 3.2)

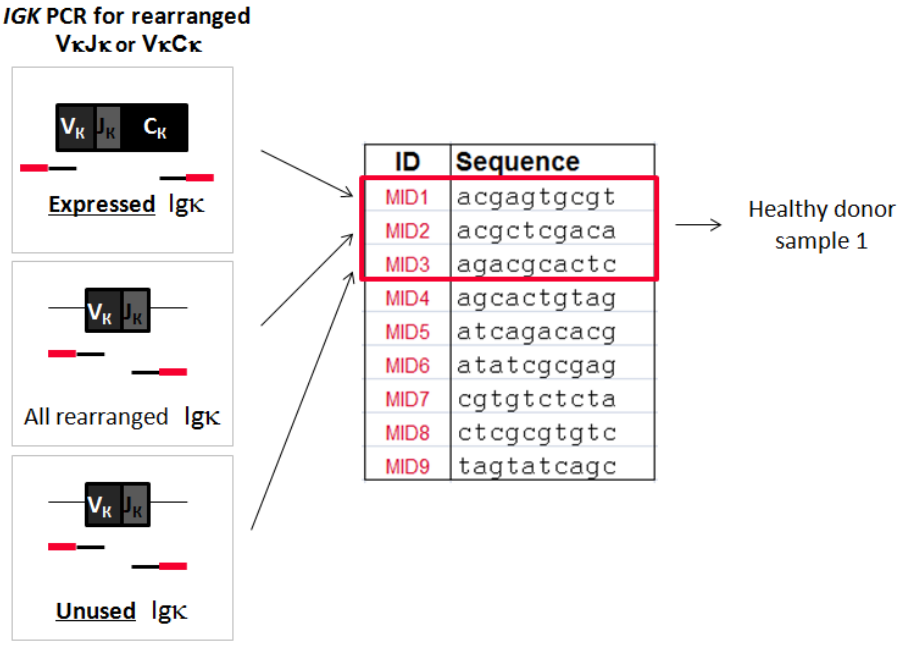


Figure 3.2 10-base multiplex-identifier (MID) tags on primers used during the second round PCR amplification of first round products. Three MID tags were assigned to each individual studied; MIDs 1-9 being used for three healthy donors and MIDs 1-9 used for three SLE donors. This was permitted since each group of MIDs 1-9 was analysed on discrete sections of the chip during the high-throughput sequencing and could therefore be identified following the analysis.

Prior to analysis, sequences were subject to quality control by eliminating sequences that were biologically implausible; this was in collaboration with Dr. Deborah Dunn-Walters

and Professor David Kipling at the University of Cardiff who designed the QC algorithm. Only sequences with an identifiable Vκ at the 5' end, followed by either a Jκ or a Cκ at the 3' end were accepted. VκCκ sequences amplified from IGK rearranged cDNA were only accepted as biologically plausible if they were over 415 nucleotides in length and VκJκ sequences amplified from the DNA of Igκ- or Igλ-expressing B cells were accepted providing they were over 290 nucleotides in length. Sequences that failed to start with a MID tag, Vκ, Jκ or Cκ were rejected.

Since the amplification method involved PCR from DNA and cDNA, it is possible that some groups of identical sequences arose from either the isolation of multiple cDNA copies of an IGK sequence from an individual cell, or amplification of a single sequence from DNA or cDNA, rather than actual clones of cells. This was addressed during analysis by identifying the CDR3; the region unique to each individual Ig gene rearrangement and critical to antigen binding, alongside the IGKV and IGKJ gene segment usage. Replicates of the same sequences were excluded so a single sequence from any "clone" was represented.

3.1.3 Experimental aims

The following table depicts the experimental conditions that were investigated for each sample, and describes the nature of each group of rearranged *IGK* sequences following the PCR amplification from either DNA or cDNA isolated populations of $CD19^+$ IgD^+ $CD27^-$ cells.

Table 3.1. Description of *IGK* sequences obtained following PCR amplification

e.g HD 1	**$CD19^+$ IgD^+CD27^-**	**DNA / cDNA**	**PCR**	**Experimental name**	**IF/OF/stop codons**	**Description**
e.g MID1	Igκ+	cDNA	Vκ–Cκ	Expressed	In-frame	Productive rearrangements of *IGK* produced & used by cell
e.g MID2	Igκ+	DNA	Vκ–Jκ	All rearranged	In-frame AND out-of-frame	Productive and non-productive rearrangements of *IGK;* some used, some unused
e.g MID3	Igλ+	DNA	Vκ–Jκ	Not expressed	In-frame AND out-of-frame	Productive and non-productive rearrangements of *IGK* **

** The groups of *IGK* rearranged sequences obtained from the Igλ+ sorted populations of B cells will contain rearrangements that are both in the correct and incorrect genetic reading frame, though both are unused by the cell since they are from Igλ–expressing cells.

Theoretically, whether the *IGK* rearrangement is genetically productive (in-frame) or non-productive (out-of-frame) it will have been functionally inactivated and therefore not transcribed and translated to protein Igκ. In-frame *IGK* rearrangements from Igλ-expressing cells will be referred to as **productively rearranged**, relating to the **genetic** status of the allele, and does **not** refer to whether the rearrangement is produced/expressed by the cell. This concept forms the framework of the chapter and investigations carried out.

The overall aim of this high-throughput sequencing analysis was to investigate the factors that shape the *IGK* repertoire of mature naive B cells in health and SLE. With reference to Table 3.1, the following questions could be addressed:

- Is there any difference between health and SLE in the imprint of the inherent light chain gene rearrangement mechanism?
- Is there any difference between health and SLE in the relative frequencies of *IGKV* gene segments in the pool of productively rearranged IGK light chains from genomic of from Igκ- and Igλ-expressing cells?
- Is there evidence of differences between health and SLE in receptor editing?
- Is there any difference between health and SLE in the expression of IGK rearrangements using different gene segments?
- Is there any difference between health and SLE in the components of the CDR3 regions of IGK light chains that are expressed compared to those that are productive forbidden and not used? Parameters investigated are stated below:
    - CDR3 length
    - GRAVY index
    - Theoretical isoelectric point (pI)
    - Aliphatic index

3.2 Results

A total of 5736 and 4698 unique IGK sequences were acquired from the pooled healthy controls and SLE respectively. These included rearrangements of *IGK* that at the genetic level were both productive and non-productive. Expressed *IGK* light chain rearrangements were those amplified from the cDNA of Igκ-producing B cells. These rearrangements were compared with *IGK* rearrangements from DNA isolated from either Igκ - or Igλ-expressing B cells, which could be further divided into two groups; those that were in the correct genetic reading frame and free from stop codons (productively rearranged), and those that were in the incorrect genetic reading frame or carried stop codons (non-productively rearranged). Productively rearranged *IGK* sequences obtained from the DNA of Igλ-expressing B cells are referred to as 'forbidden productive'. A summary of the frequency of *IGK* rearrangements pooled for each healthy donor and SLE sample obtained is shown in Figure 3.3 A. The number of sequences obtained from each individual can be found in Appendix 1.

Forbidden productive rearrangements of *IGK* are in the correct genetic reading frame and therefore productive at the genetic level, but since they are present in Igλ-producing cells cell fate decisions must have been made to not express them.

Forbidden productive rearrangements are likely to be the consequence of an aspect of functionality relating to the specificity of the heavy/light chain pairing or inability to make a functional receptor with the heavy chain. The frequency of forbidden productive rearrangements from Igλ-expressing mature naive B cells was equivalent in health and SLE (Figure 3.3 B).

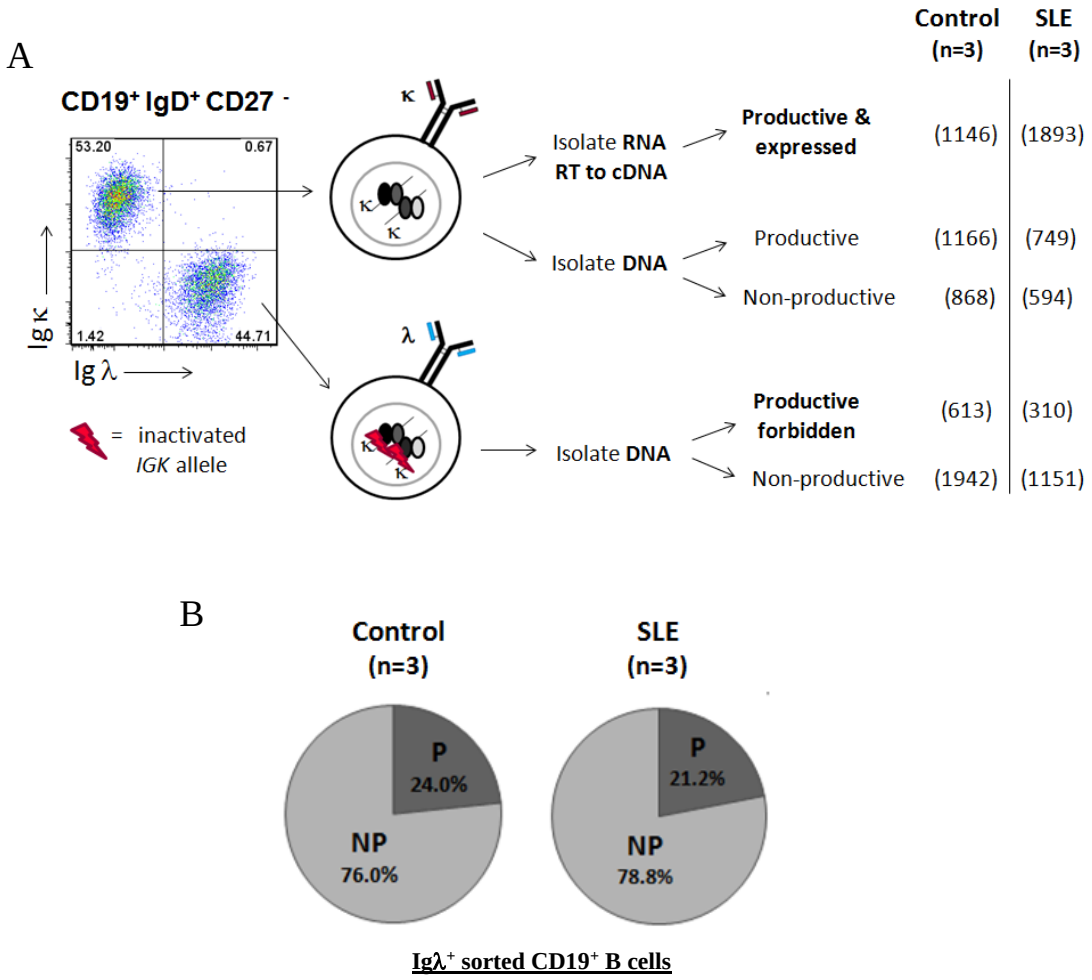


Figure 3.3 High throughput sequence analysis. (A) Frequencies of unique, QC-passed rearranged sequences following high-throughput sequencing of the *IGK* light chain repertoire. Schematic outlines the frequency of *IGK* rearranged sequences obtained from pooled healthy control and pooled SLE FACS-sorted Igκ- and Igλ-expressing mature naive B cells after high-throughput sequencing analysis. **(B)** Percentage of forbidden productive *IGK* rearrangements sequenced from genomic DNA isolated from Igλ-expressing mature naive B cells. Forbidden productive rearrangements were present at equivalent frequencies in both health (24%) and SLE (21.2%).

3.2.1 Analysis of *IGKV* and *IGKJ* segment rearrangement frequencies

Biases in *IGKV* gene segment usage have been observed in SLE, but their origin and significance are unknown. This section investigates the origin of these biases by studying the profile of *IGKV* gene rearrangements in productively and non-productively rearranged genomic DNA from Igκ- and Igλ-expressing mature naive B cells. These *IGK* rearrangement profiles will be compared with the *IGKV* gene segment rearrangement frequency of the expressed *IGK* repertoire, determined through studying the cDNA of Igκ-expressing cells.

3.2.1.1 Study of unused *IGK* rearrangements

Studying unused *IGK* sequences allows investigation into the processes underlying *IGK* gene segment expression and selection. Unused sequences were analysed from genomic DNA and included non-productively rearranged alleles of *IGK* from Igκ- and Igλ-expressing B cells, and the forbidden productive *IGK* rearrangements from Igλ-expressing B cells. No significant differences were observed in the non-productive rearrangements from Igκ-expressing cells when rearrangements from healthy donors and patients with SLE were compared (Figure 3.4 A). Biases towards rearrangement of two infrequently expressed V gene segments *IGKV3D-20* ($p=4.2x10^{-7}$) and *IGKV2-24* ($p=1.78x10^{-2}$) were observed in SLE compared to health in the non-productively rearranged IGK sequences from Igλ-expressing cells. (Figure 3.4 B). Forbidden productive rearrangements of *IGK* appeared to have a different profile to the two sets of non-productively rearranged *IGK* sequences, though the only difference observed between health and SLE was a bias towards rearrangement of the gene segment *IGKV3D-20* ($p=7.33x10^{-3}$) in SLE which is rarely involved in rearrangement events. (Figure 3.4 C).

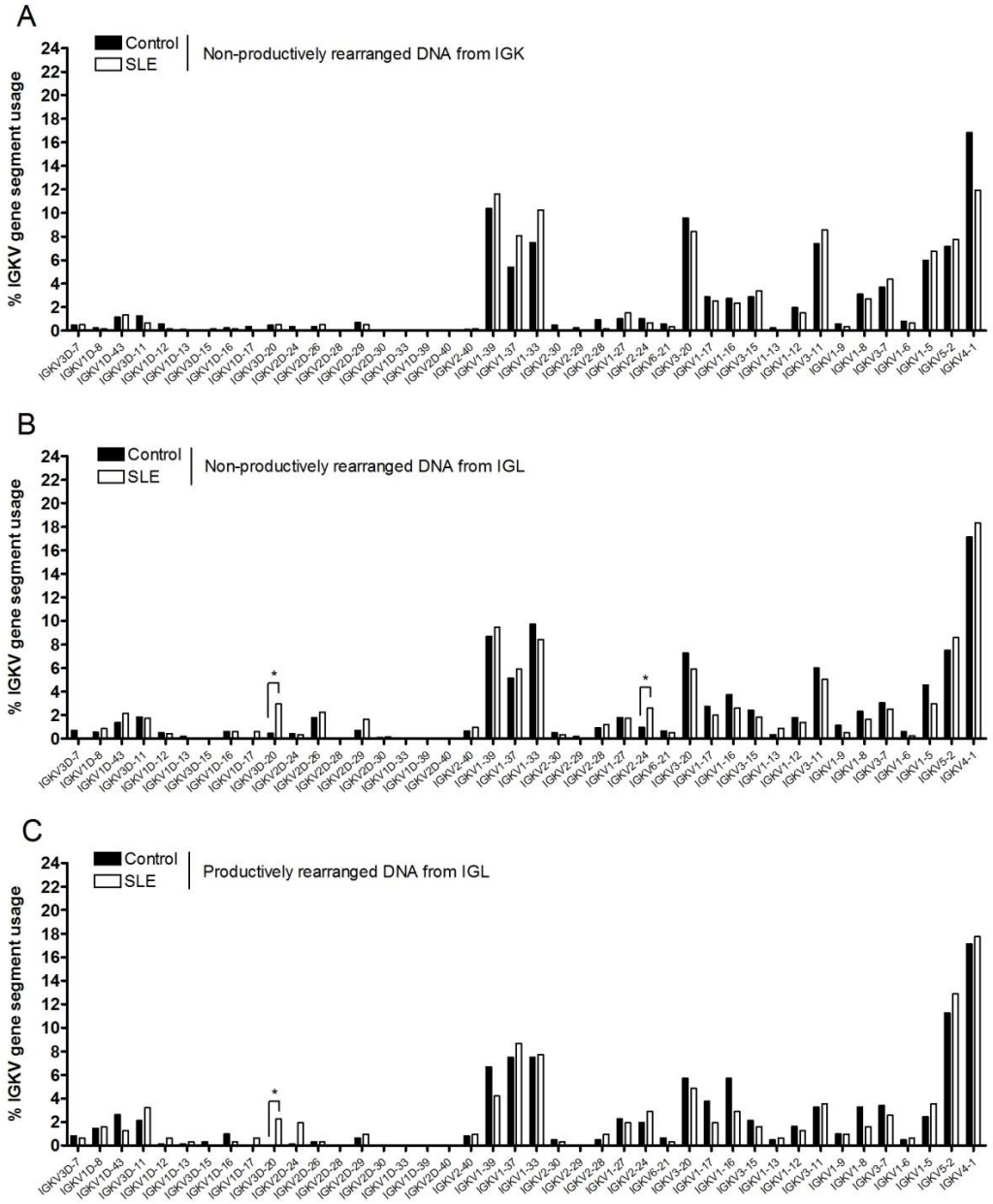


Figure 3.4 Relative rearrangement frequency of *IGKV* gene segments in the unused *IGK* repertoire of mature naive B cells in health (n=3) and SLE (n=3). (A) Non-productive *IGK* rearrangements from Igκ-expressing B cells. (B) Non-productive *IGK* rearrangements from Igλ-expressing B cells where biases towards rearrangement of two *IGKV3D-20* and *IGKV2-24* were observed in SLE compared to health and (C) Forbidden productive *IGK* repertoire from Igλ -expressing cells shows an SLE bias towards rearrangement of the gene segment *IGKV3D-20*. Statistical analysis was performed using χ^2 analyses and adjusted by Bonferroni correction (*).

This section demonstrates that selection of *IGKV* gene segments and the inherent biases in the gene rearrangement process at the DNA level as the repertoire develops is equivalent in health and SLE.

3.2.1.2 Comparison of productive *IGK* rearrangements in genomic DNA from Igκ- and Igλ-expressing cells

It was considered that any productively rearranged *IGKV* segments that were present with greater frequency in the forbidden productive sequences from Igλ-expressing cells than the productively rearranged group from Igκ-expressing cells had been selected against based on an aspect of functionality during B cell development. Significant selection against inclusion of the *IGKV5-2* ($p=1.7x10^{-15}$), *IGKV1-37* ($p=7.69x10^{-7}$), *IGKV1-16* ($p=1.92x10^{-7}$), *IGKV3-7* ($p=1.44x10^{-2}$), and the infrequently rearranged *IGKV* gene segments *IGKV1D-43* ($6.37x10^{-4}$), *IGKV3D-11* ($9.74x10^{-3}$) and *IGKV1D-16* ($p=2.79x10^{-2}$) into the rearranged *IGK* repertoire in genomic DNA of Igκ-expressing cells was observed in health (in order of gene segment frequency) (Figure 3.5 A). Selection against *IGKV5-2* ($p=3.24x10^{-10}$), *IGKV1-37* ($p=1.02x10^{-3}$), *IGKV2-24* ($p=1.1x10^{-4}$) and *IGKV3D-11* ($p=2.46x10^{-4}$) was also apparent in genomic DNA in SLE (Figure 3.5 B).

In addition, it was considered that any *IGKV* segments that were present with lower frequency in the forbidden productive sequences than in the productively rearranged group from Igκ-expressing cells had been selected *for* based on an aspect of functionality during B cell development. Selection for inclusion of *IGKV1-39* ($p=2.0x10^{-5}$), *IGKV3-20* ($p=6.7x10^{-6}$), *IGKV3-11* ($1.32x10^{-9}$), *IGKV1-5* ($p=5.2x10^{-6}$) and *IGKV3-15* ($p=4.45x10^{-2}$) segments into the repertoire in genomic DNA of Igκ-expressing cells was observed in

health (in order of gene segment frequency) (Figure 3.5 A). With the exception of *IGKV1-5* where selection for this segment was apparent as a trend, all of these *V* segments were also significantly selected for in SLE. Inclusion of *IGKV1-39* ($p=2.84x10^{-4}$), *IGKV3-20* ($p=8.15x10^{-5}$), *IGKV3-11* ($9.52x10^{-4}$) and *IGKV3-15* ($p=2.84x10^{-2}$) SLE segments was observed (Figure 3.5 B).

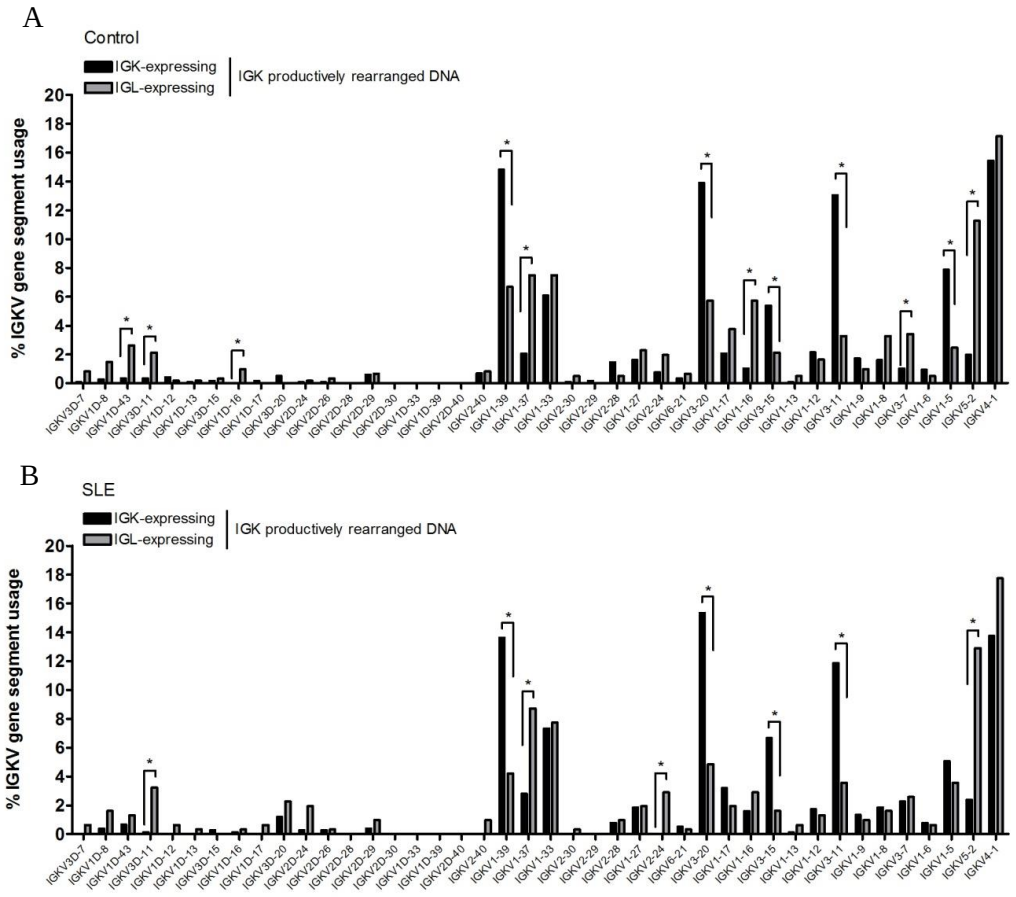


Figure 3.5 Comparison of the relative rearrangement frequency of *IGKV* gene segments in the productive *IGK* rearranged repertoires of mature naive B cells from the genomic DNA of Igκ- and Igλ -expressing cells in (A) healthy controls (n=3) and (B) SLE (n=3). Evidence of significant *IGKV* segment selection for and against expression was observed and were mirrored by SLE as observed in health, with the exception of segment *IGKV1-5*. Statistical analysis was performed using χ^2 analyses and adjusted by Bonferroni correction. (*)

If the assumptions made regarding *IGKV* segment selection were correct, then the frequencies of non-productive *IGK* rearrangements from Igκ- and Igλ-expressing cells that have no biases should not only be equal to each other, but should also lie between the frequencies of productive rearrangements in Igκ-expressing cells and the forbidden productive rearrangements of *IGK* in Igλ-expressing cells where selection occurred.

Consistent with this interpretation, where selection either for or against a particular *IGKV* gene segment was identified, the non-productive rearrangements were present at relative frequencies between those of productive repertoire from Igκ-expressing cells and the forbidden productive sequences in health and SLE (Figure 3.6).

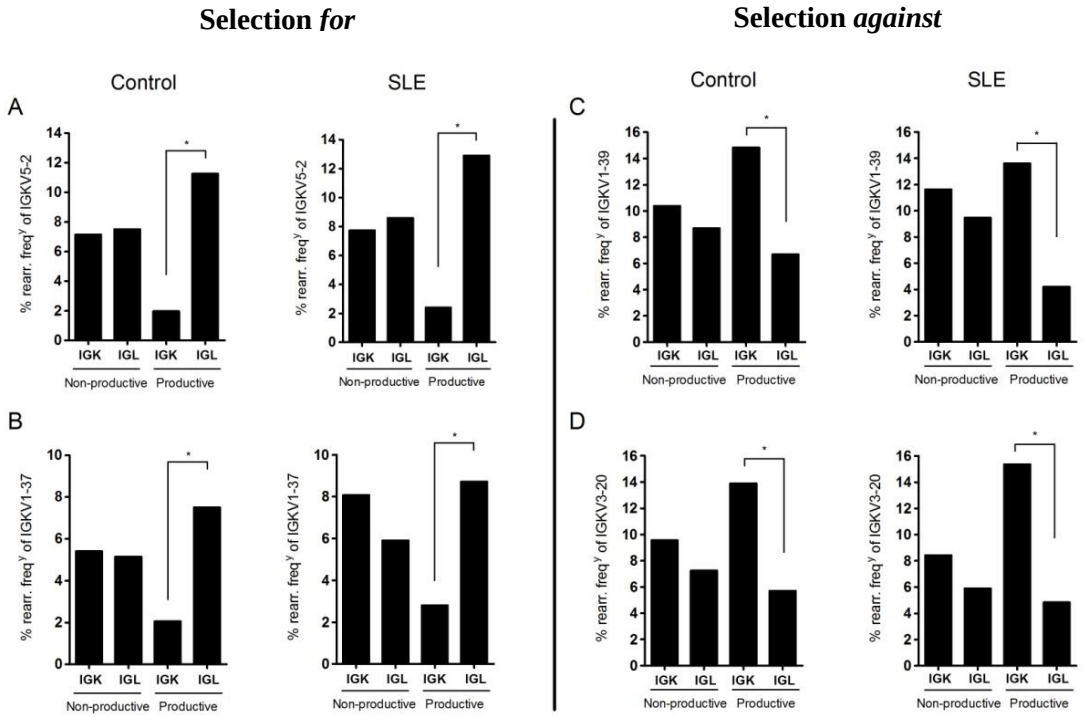


Figure 3.6 Evidence for gene selection for and against *IGKV* segments in health and SLE. (A and B) Selection against segments *IGKV5-2* and *IGKV1-37* into the productive IGK repertoire. (C and D) Selection for segments *IGKV1-39* and *IGKV3-20* into the productive *IGK* mature naive repertoire. Relative frequencies of V segments in non-productive *IGK* rearrangements from genomic DNA of Igκ- and Igλ-expressing cells lie between the relative frequencies of productive rearrangements from Igκ-expressing cells and productive forbidden productive *IGK* rearrangements from Igλ-expressing cells using these specific *IGKV* segments Statistical analyses performed using χ^2 analyses and adjusted by Bonferroni correction. (*)

3.2.1.3 Comparison of expressed *IGK* rearranged sequences from cDNA of Igκ-expressing cells with productive *IGK* rearrangements from the genomic DNA of *IGK*-expressing cells

The *IGKV* segment repertoire of expressed sequences from the cDNA of Igκ-expressing cells was analysed in health and SLE. The two most 3' variable gene segments *IGKV4-1* ($p=3.74x10^{-10}$), and *IGKV5-2* ($p=5.00x10^{-2}$) were significantly more frequently observed in the expressed repertoire of mature naive B cells from patients with SLE compared to controls (Figure 3.7 A) There was also a significantly lower frequency of rearrangements involving *IGKV3-11* ($p=5.99x10^{-4}$), and *IGKV3-15* ($p=7.42x10^{-3s}$) gene segments in the expressed repertoire in SLE compared to health. These biases were not observed in the productive rearrangements from genomic DNA from the same isolated cells from the same individuals (Figure 3.7 B).

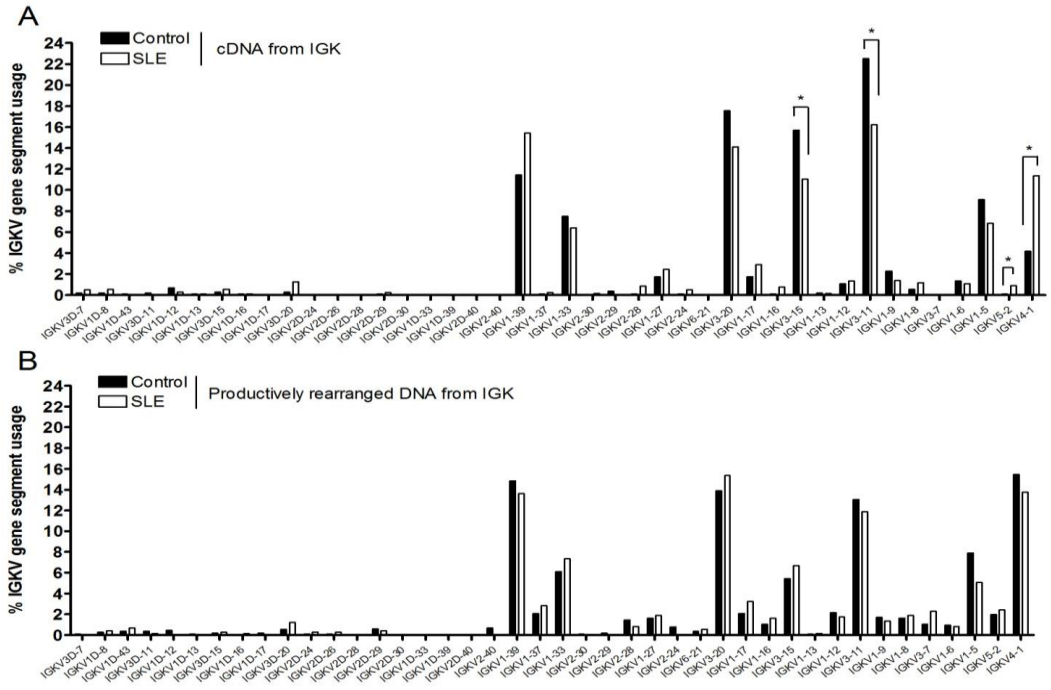


Figure 3.7 *IGKV* gene segment bias in mature naive B cells in SLE is apparent in the expressed but not the rearranged *IGK* repertoire. **(A)** Relative frequency of *IGKV* gene segment expression where *IGKV4-1* and *IGKV5-2* were significantly more frequently observed in the expressed repertoire in SLE compared to controls, and *IGKV3-11* and *IGKV3-15* significantly less frequently observed in SLE compared to health. **(B)** Relative frequency of *IGKV* gene segment usage in productive rearrangements of *IGK* amplified from DNA from FACS-sorted Igκ-expressing B cells in controls (n=3) and SLE (n=3). These biases were not observed in the productive rearrangements from gDNA from the same isolated cells from the same individuals. Statistical analyses were performed using χ^2 analyses and adjusted by Bonferroni correction (*).

3.2.1.4 Comparison of the relative rearrangement frequencies of gene segments *IGKV4-1* and *IGKV5-2* in *IGK* sequences that were expressed, productively rearranged in the genomic DNA from Igκ-expressers and forbidden productive rearrangements.

Gene segments *IGKV4-1* and *IGKV5-2* were of particular interest because the bias observed was towards preferential expression in SLE compared to health. In addition, a bias towards *IGKV4-1* expression in SLE was a consistent finding in previous studies (Andrews et al., 2013). The segment *IGKV5-2* was stringently selected against during development because the productive rearrangements in Igκ-expressing cells were less frequent than forbidden productive rearrangements using *IGKV5-2* in Igλ-expressing cells in both health and SLE ($p=1.52x10^{-15}$). This segment was further selected against at the level of expression in health ($p=6.19x10^{-5}$) but not in SLE (Figure 3.8 A). The *IGKV4-1* segment was not selected against during development because the productive rearrangements in Igκ-expressing cells were not significantly less frequent than in the forbidden productive rearrangements in Igλ-expressing cells in both health and SLE. However, this segment was selected against at the level of expression compared to the productive rearrangements of *IGK* from Igκ-expressing B cells ($p=4.15x10^{-18}$) and the forbidden productive rearrangements from Igλ-expressing B cells ($p=1.56x10^{-18}$) in health, whereas this was not observed in SLE (Figure 3.8 B).

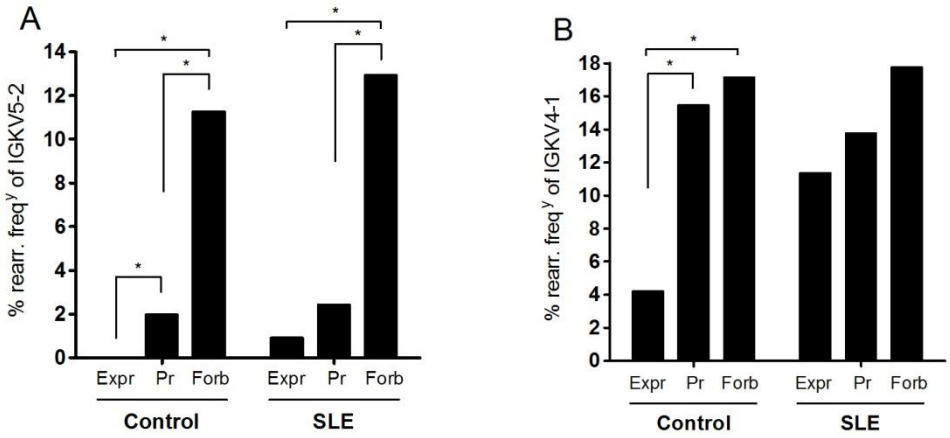


Figure 3.8 Evidence to demonstrate the selection processes of *IGK* rearrangements involving individual *IGKV* gene segments. **(A)** Selection against *IGKV5-2* at the DNA level and at the level of gene expression in both health (n=3) and SLE (n=3). **(B)** Lack of selection against *IGKV4-1* at the DNA level in health and SLE, but selection against expression of *IGKV4-1* in health. Failure to select against expression of *IGKV4*-1 in SLE compared to health. Statistical analyses were performed using χ^2 analyses and adjusted by Bonferroni correction (*).

These data provide evidence to suggest that Ig light chain expression could be regulated by more than one mechanism; at the molecular level during the initial Ig gene rearrangement process, as suggested by the usage of *IGKV5-2* in both the healthy and SLE B cell repertoire, and also at the level of expression. Usage of *IGKV4-1* appears to be selected against only at the level of expression of healthy subjects, where this is not observed in SLE.

3.2.1.5 Analysis of *IGKJ* segment usage in genomic DNA

It has been previously described that usage of more distal *IGKJ* gene segments may reflect receptor editing (Bräuninger et al., 2001). To further investigate the factors that shape the *IGK* expressed repertoire that may indicate receptor editing and whether this is the same in health and SLE, *IGKJ* gene segment rearrangement frequency was compared in *IGK*

sequences obtained from productively and non-productively rearranged genomic DNA in health and SLE. No selection for or against *IGKJ* segments was observed when the productive rearrangements of *IGK* from Igκ-expressing cells were compared with the forbidden productive *IGK* rearrangements from Igλ-expressing cells. The only significant observation was a low frequency of non-productive rearrangements of *IGKJ4* that was apparent in both health and SLE. No differences were found in the relative frequency of *IGKJ* segment involvement in health or SLE, and did not appear to be influenced by whether the *IGK* rearrangements were productive or non-productive (Figure 3.9).

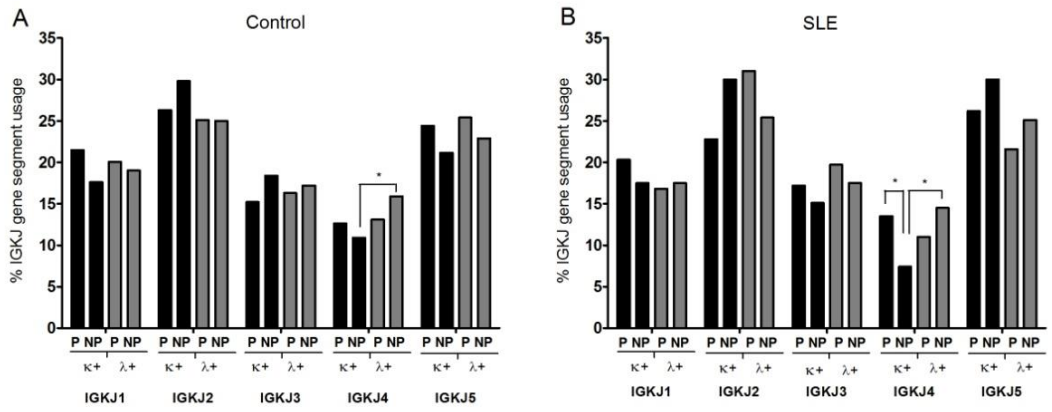


Figure 3.9 *IGKJ* gene segment rearrangement frequency in productive and non-productive rearrangements from both Igκ- and Igλ-expressing B cells in **(A)** controls and **(B)** SLE. No selection for or against *IGKJ* segments was observed when the productive rearrangements of *IGK* from Igκ-expressing cells were compared with the forbidden productive *IGK* rearrangements from Igλ-expressing cells, though Statistical analysis was performed using χ^2 analyses and adjusted by Bonferroni correction (*).

This section provides evidence for equivalent repertoire usage of the *IGKJ* gene segments in health and SLE.

3.2.2 Comparison of the features of the CDR3 by expressed *IGK* rearrangements from the cDNA of Igκ-expressing cells with forbidden productive rearrangements of *IGK* from Igλ-expressing B cells.

The hallmark of SLE is the production of antibodies specific for self-antigen, so it is of interest to study the components of the rearrangement that encodes a portion of the antigen binding site. The CDR3 region is associated antigen binding; it forms a portion of the antigen binding site in the folded receptor. Since the CDR3 regions tend to be the most hypervariable and therefore have the most influence over antigen specificity, the properties of *IGK* CDR3s that were expressed by Igκ-producing B cells were analysed and compared to the CDR3s of forbidden productive *IGK* rearrangements from Igλ-expressing cells in health and SLE. The CDR3s of Ig light chains range from approximately 8 to 11 amino acids in length so are significantly shorter than the CDR3s of the heavy chains, but it was considered important to investigate any variability in the light chain CDR3s since they too contribute to the antigen specificity of the antibody. The CDR3 properties analysed are stated below and the reasons for studying these parameters will be discussed as a prelude to each data section:

- Analysis of CDR3 length
- GRAVY index analysis of the CDR3
- Analysis of the theoretical isoelectric point (pI) of the CDR3
- Analysis of the aliphatic index of the CDR3

Non-productively rearranged *IGK* sequences obtained from the genomic DNA of both Igκ- and Igλ-expressing cells were not considered as they are either out-of-frame or contain stop codons and are consequently not produced.

3.2.2.1 Analysis of CDR3 length

Longer CDR3 lengths in Ig heavy chains have previously been associated with DNA-binding antibodies (Aguilera et al., 2001; Crouzier et al., 1995; Radic and Weigert, 1995). Shorter CDR3 lengths are associated with antigenic selection after somatic hypermutation and are a feature of switched memory B cells (Wu et al., 2010).

The CDR3 length of expressed *IGK* rearrangements from the cDNA of Igκ-expressing cells in health and SLE was first investigated, and compared to forbidden productive *IGK* rearrangements from the genomic DNA from Igλ-expressing cells. In the expressed *IGK* repertoire, a significantly higher proportion of *IGK* sequences in SLE ($p=9.22x10^{-3}$) expressed Igκ light chains with CDR3 regions 9 amino acid residues in length compared to health (Figure 3.10). Despite this however, there was a tendency to express rearranged *IGK* alleles with CDR3s 9 residues in length in both health and SLE (Figure 3.10). Forbidden productive rearranged *IGK* CDR3 regions from Igλ-expressing B cells tended to be longer than those observed in the expressed naive repertoires in health and SLE. A significantly higher proportion of forbidden productive *IGK* rearrangements had CDR3s 10 residues long compared to those in the expressed repertoires in both health ($p=2.01x10^{-7}$) and SLE ($p=2.35x10^{-8}$). There was also a significantly higher proportion of forbidden productive *IGK* rearrangements with CDR3 lengths outside the range of 8-11 residues in length than those observed in the expressed IGK repertoires in health ($p=1.54x10^{-8}$) and SLE ($p=3.43x10^{-7}$).

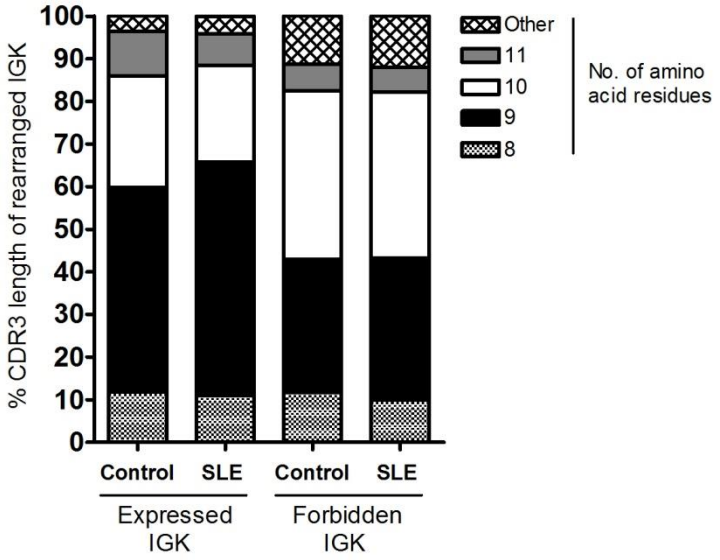


Figure 3.10 CDR3 length variability in productively rearranged *IGK* sequences that were expressed compared with those that were forbidden productive in healthy controls and SLE. Statistical analyses were performed by two-tailed χ^2 analyses and adjusted by Bonferroni correction.

This section suggests that shorter CDR3 lengths are associated with the expressed *IGK* repertoire in health and SLE, compared to forbidden productive IGK rearrangements that have been selected against during B cell development for unknown reasons. A significantly larger proportion of expressed *IGK* sequences with shorter CDR3s (8-9 residues) were observed in SLE compared to health.

3.2.2.2 GRAVY index analyses

The grand average of hydropathicity (GRAVY) index is a score that can be used to identify the solubility of a protein based on its amino acid composition. It is calculated as the sum of individual amino acid hydropathy values divided by the number of residues in the total CDR3 (Kyte and Doolittle, 1982). A negative GRAVY index indicates a hydrophilic protein, whereas a positive score would be given to a hydrophobic protein. Due to the

varying lengths of CDR3 regions after *IGK* rearrangements and how this would influence the GRAVY score assigned, sequences were analysed according to their total CDR3 length of 8, 9, 10 or 11 amino acid residues.

Differences in the hydropathicity of the CDR3 regions could be observed between *IGK* rearrangements in health that were expressed, compared to those that were productively rearranged but forbidden. All CDR3 regions of *IGK* rearranged light chains were hydrophilic, though there was a tendency for forbidden productive *IGK* rearrangements to have less negative GRAVY indices than those that are expressed in health. This was apparent in CDR3 regions consisting of all lengths (Figure 3.11 A-D). A similar trend is displayed by comparing the expressed and forbidden productive *IGK* rearranged sequences in SLE, where all CDR3 regions are hydrophilic, but forbidden productive *IGK* rearrangements tended to be less hydrophilic. This was significant in CDR3 regions of 9-11 amino acid residues in length, but this was only observed as a trend in CDR3 regions 8 residues long (Figure 3.11 E-H).

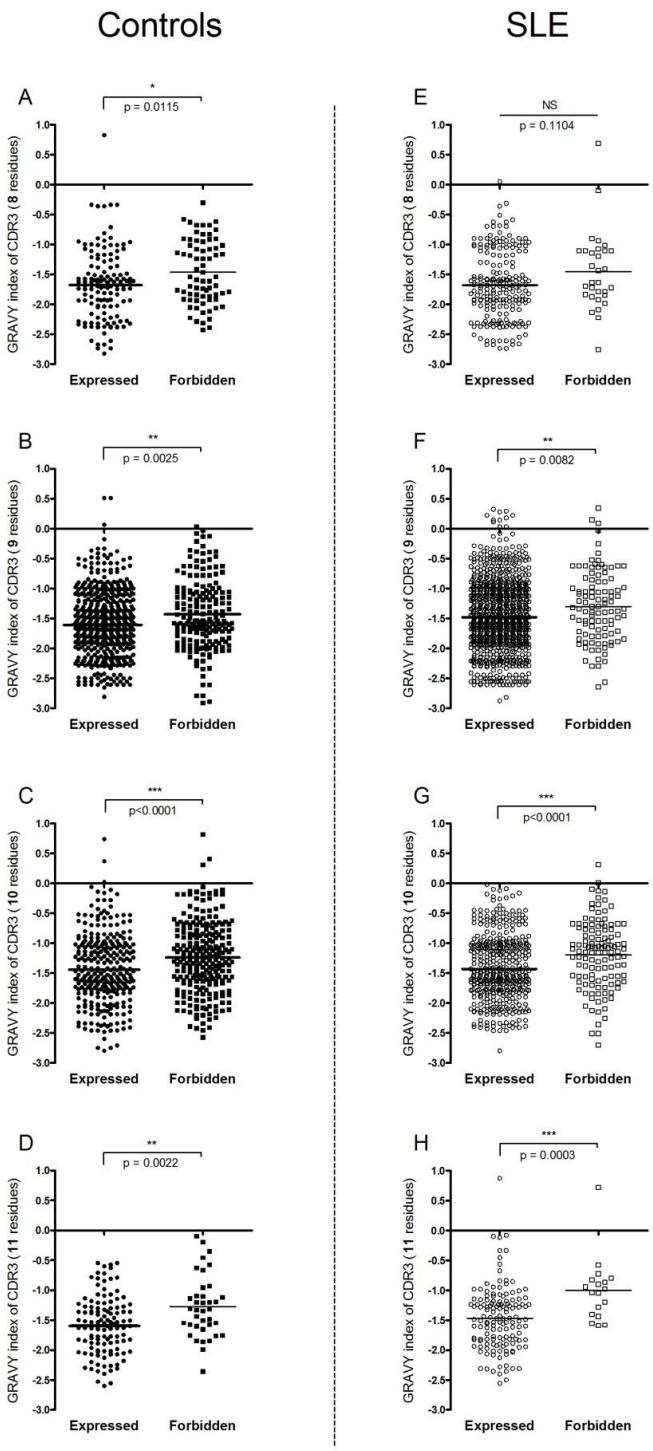


Figure 3.11 GRAVY index analysis of CDR3 regions of expressed and forbidden productive *IGK* rearrangements of varying CDR3 length in **(A-D)** healthy controls (n=3) and **(E-H)** SLE patients (n=3). Statistical analysis was determined a two-tailed Student's t-test assuming unequal variance. * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

GRAVY index scores among the expressed and forbidden productive *IGK* repertoires were then compared between health and SLE. CDR3 regions of 8, 10 and 11 amino acid residues in length did not differ in their hydropathicity in health or SLE (Figure 3.12 A, C and D respectively.) Expressed *IGK* light chains of 9 residues in length however not only appear to be the most frequently expressed CDR3 length, but those expressed in SLE tend to be less hydrophilic, with a less negative GRAVY score compared to healthy controls (Figure 3.12 B.) Comparative analysis of the forbidden productive CDR3 regions within the control and SLE *IGK* repertoires revealed that regardless of whether CDR3 regions were 8, 9, 10 or 11 amino acid residues in length, hydropathicity scores assigned were of a similar hydropathicity, tending to be hydrophilic in nature in health and SLE (Figure 3.12 E-H).

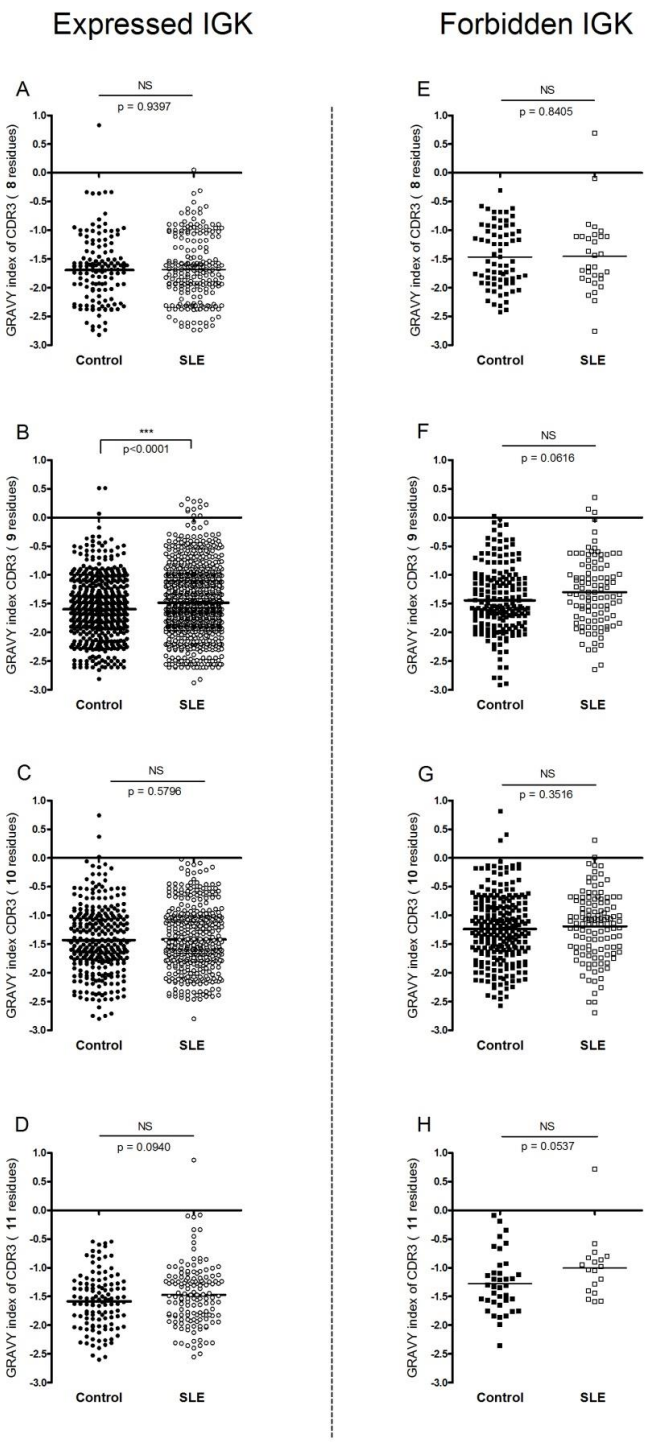


Figure 3.12 Analysis of GRAVY index scores assigned to CDR3 regions of varying amino acid length comparing productively rearranged *IGK* sequences in health (n=3) and SLE (n=3). **(A-D)** The expressed *IGK* repertoire and **(E-H)** the forbidden productive *IGK* repertoire. Statistical analysis was determined a two-tailed Student's t-test assuming unequal variance. *p<0.05, **p<0.01, *** p<0.001.

This section demonstrates that in both health and SLE *IGK* light chain CDR3s are hydrophilic, but rearrangements of *IGK* that are expressed tend to be more hydrophilic than the forbidden productive repertoire of *IGK* rearrangements. This appears to be independent of total CDR3 length. Comparing the expressed *IGK* repertoires in health and SLE, CDR3s 9 amino acid residues in length are less hydrophilic in SLE than those of the same length in health. No differences in hydropathicity were observed within the forbidden productive repertoire of any CDR3 length studied between health and SLE.

3.2.2.3 Analysis of theoretical isoelectric points of *IGK* CDR3s in health and SLE

The isoelectric point (pI) is the pH at which the net charge of a protein is zero. Immunoglobulins are highly diverse in their isoelectric points in that they can fall within a range of pH 4 to 9. A protein will carry a net positive charge if the pH of its environment is less than its pI, and a net negative charge at a pH above its isoelectric point. The theoretical pI of the CDR3 regions in expressed and forbidden productive *IGK* rearranged sequences was analysed in healthy controls. Analyses revealed that CDR3 regions from forbidden productive *IGK* rearrangements that were 10 and 11 amino acid residues in length tended to have significantly lower theoretical pI values that those *IGK* rearrangements that were expressed (Figure 3.13 C and D). No differences in pI were observed in CDR3 regions 8 and 9 residues long (Figure 3.13 A and B.) Similar trends in the pI of CDR3 regions were found in SLE through comparing *IGK* sequences that were either expressed or forbidden productive. Regardless of whether the *IGK* sequences were expressed or in the forbidden productive repertoire, rearrangements with CDR3 regions of 8, 9 or 10 amino acids in length on average had similar pI scores (Figure 3.13 E, F and G). A significantly lower pI was found in the forbidden productive *IGK* sequences in SLE when CDR3 regions were 11 amino acids in length (Fig. 3.13 H).

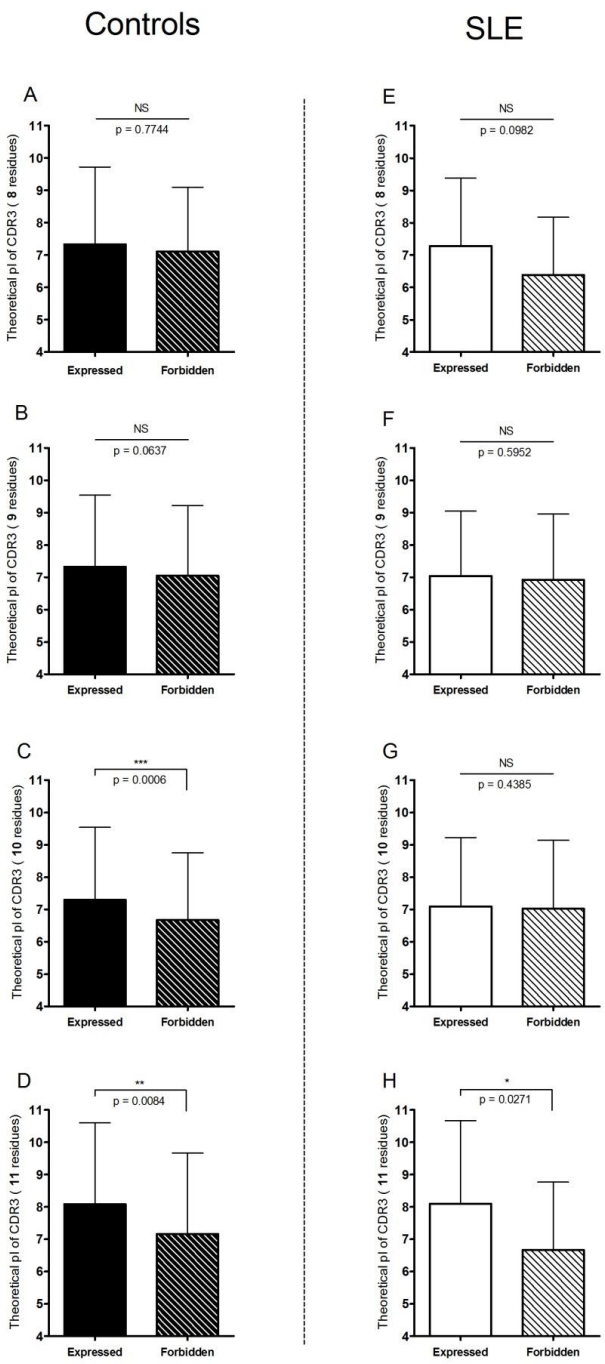


Figure 3.13 Analysis of the theoretical isoelectric point (pI) of CDR3 regions of expressed and forbidden productive *IGK* rearrangements of varying CDR3 length in **(A-D)** healthy controls (n=3) and **(E-H)** patients with SLE (n=3). Statistical analysis was determined a two-tailed Student's t-test assuming unequal variance Error bars are SD. * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

Theoretical pI scores among the expressed and forbidden productive *IGK* repertoires were then compared between health and SLE. Comparing health with SLE, theoretical pI scores tended to be similar regardless of CDR3 length when studying the expressed *IGK* repertoire, though a significantly lower pI was observed when analysing expressed CDR3 lengths of 9 amino acid residues long in SLE. No other differences were identified (Figure 3.14 A-H).

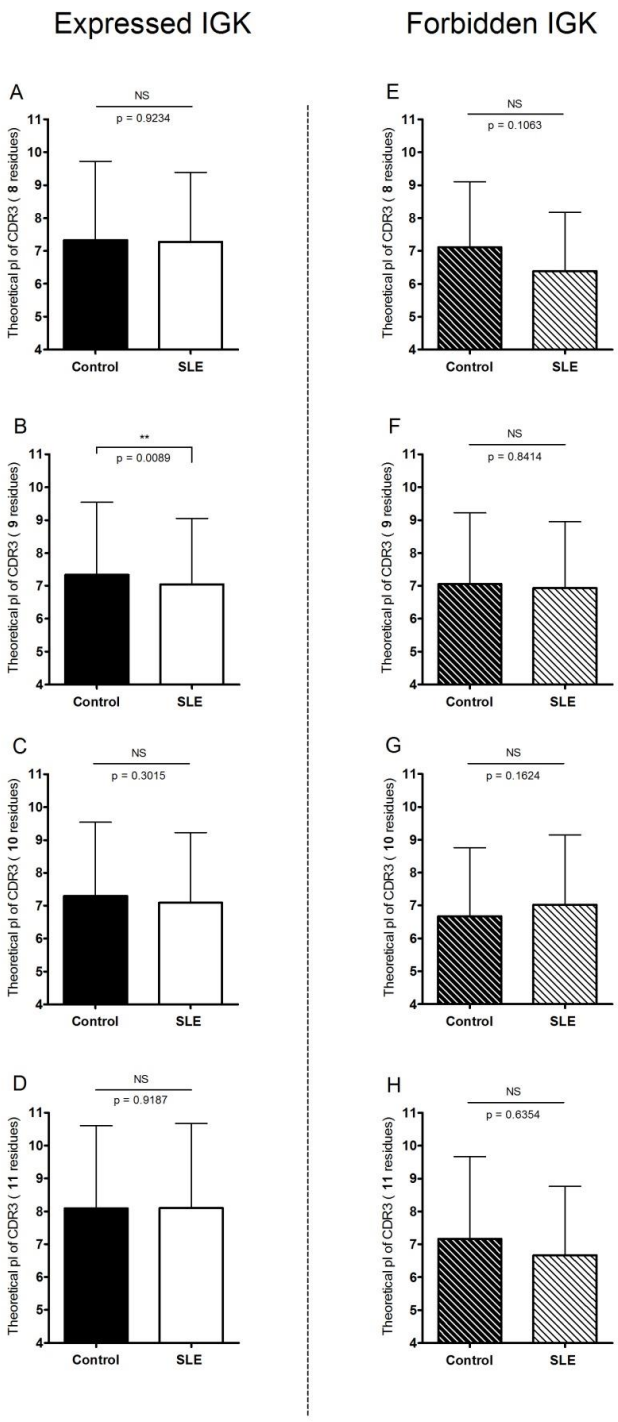


Figure 3.14 Analysis of the theoretical isoelectric point (pI) of CDR3 regions of expressed and forbidden productive in-frame IGK rearrangements of varying CDR3 length in **(A-D)** expressed IGK repertoire and **(E-H)** forbidden productive IGK repertoire in health and SLE. Statistical analysis was determined a two-tailed Student's t-test assuming unequal variance. Error bars are SD. * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

This section illustrates that forbidden productive rearranged *IGK* sequences with CDR3 lengths of 10 or 11 amino acids in length tend to have a lower pI compared to those that were expressed in health. SLE CDR3s 11 residues long also had lower theoretical pIs in the forbidden productive *IGK* repertoire compared to the expressed. In addition, the pI of the expressed *IGK* repertoire 10 residues in length tended to be greater in health than SLE.

3.2.2.4 Analysis of aliphatic indices of *IGK* CDR3s in health and SLE

Further insight into the biochemical properties of rearranged *IGK* CDR3 regions can be gained by studying the aliphatic index of the CDR3, which relates to the hydrophobicity of a globular protein through determination of the relative volume occupied by aliphatic side chains (valine, alanine, isoleucine and leucine.) Higher aliphatic indices are considered to indicate globular proteins that are more thermostable and less polyspecific than those with lower aliphatic indices. These with lower aliphatic indices are associated with a higher tendency to exhibit polyspecificity.

High-throughput sequence analysis revealed that productive rearrangements of *IGK* of 9, 10 and 11 amino acid residues in length have a significantly lower aliphatic index if they are expressed, compared to those that are forbidden productive and not expressed by the cell. (Figure 3.15 A-D) This observation was the same when comparing the expressed and forbidden productive *IGK* repertoires in SLE, where forbidden productive sequences of 9, 10 and 11 residues tended to have a significantly higher aliphatic index than those that were expressed. (Figure 3.15 E-H). No differences in aliphatic index were observed in CDR3 regions 8 residues long.

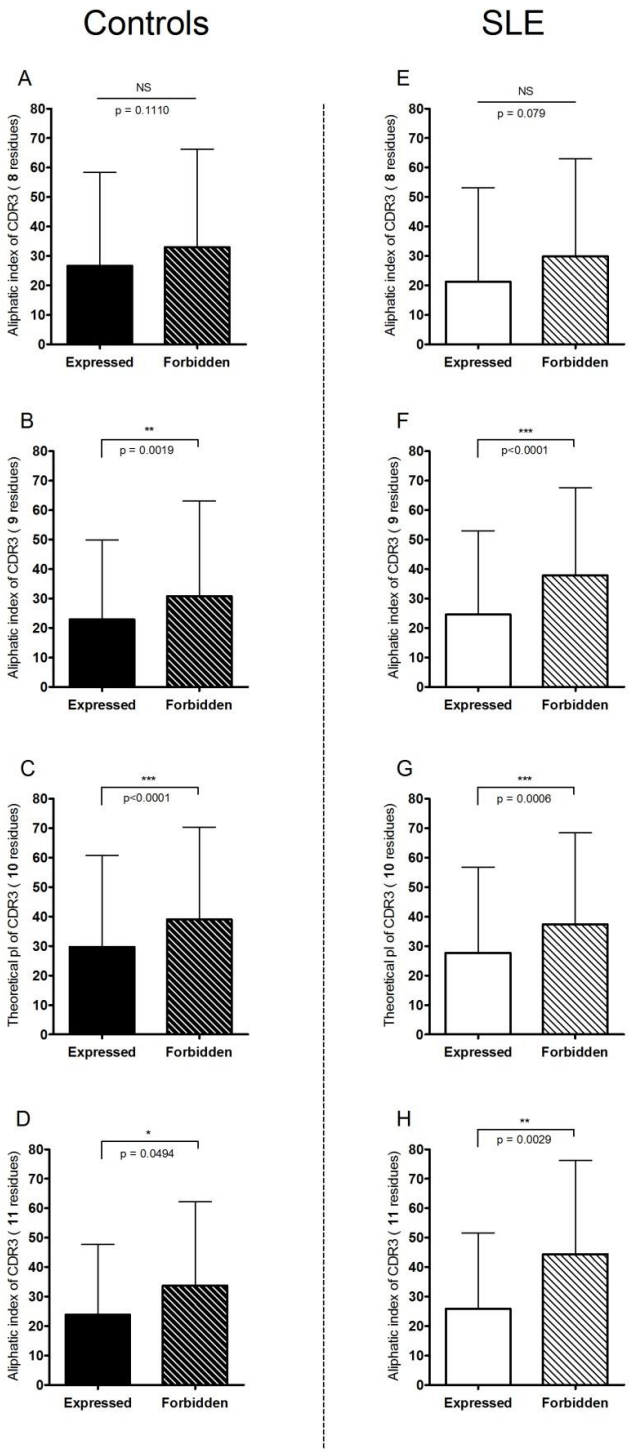


Figure 3.15 High-throughput sequencing analysis to determine the aliphatic index of CDR3 regions of in-frame *IGK* expressed and forbidden productive rearrangements of increasing CDR3 length in **(A-D)** controls (n=3) and **(E-H)** SLE (n=3). Statistical analysis performed by student's t-test. *p<0.05, **p<0.01, *** p<0.001.

Comparison of aliphatic indices in expressed *IGK* repertoires between health and SLE indicated that CDR3 regions 9 amino acid residues in length were significantly lower in SLE than health, though no differences were observed in *IGK* sequences of other CDR3 lengths studied (Figure 3.16 A-D). Within the forbidden productive repertoires of *IGK* sequences, no differences in aliphatic index were observed between health and SLE (Figure 3.16 E-H).

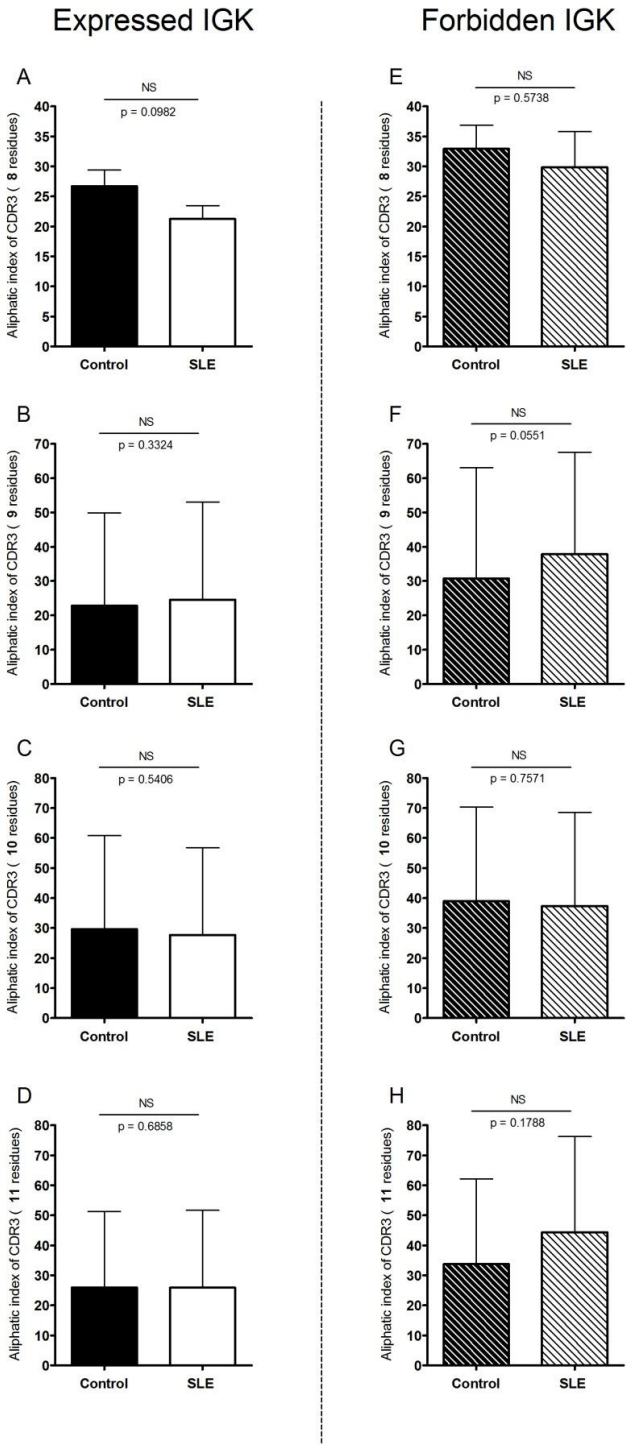


Figure 3.16 Analysis of the aliphatic index of CDR3 regions of varying amino acid length comparing in-frame IGK rearrangements in health (n=3) and SLE (n=3) within (**A-D)** the expressed IGK repertoire and **(E-H)** the forbidden productive IGK repertoire. Statistical analysis was determined a two-tailed Student's t-test assuming unequal variance. * $p<0.05$, **$p<0.01$, *** $p<0.001$.

This section shows that in both health and SLE, the aliphatic index of the *IGK* CDR3 is significantly lower in the expressed repertoire, compared to the inactivated repertoire in CDR3s of 9-11 amino acid long. Furthermore, expressed IGK CDR3s 9 residues long in SLE tended to have lower aliphatic indices than those *IGK* rearrangements expressed in health.

3.2.2.5 Summary of experimental findings

Below is a summary of findings and interpretations of the origins of *IGK* light chain biases with reference to *IGKV* and *IGKJ* gene segment frequency of rearrangement in the productive and non-productive groups of *IGK* rearranged sequences analysed from both the Igκ- and Igλ-expressing mature naive B cell repertoires in health and SLE after high-throughput sequencing analysis.

- **Is there any difference between health and SLE in the imprint of the basic rearrangement mechanism?**

  Outcome: The two data sets of non-productive sequences were virtually identical between Igκ- and Igλ-expressing cells as a source of the sequences in health and SLE, demonstrating no difference between the rearrangement processes in these groups.

- **Is there any difference between health and SLE in the relative frequencies of IGKV gene segments in the pool of productively rearranged IGK light chains in genomic DNA from Igκ- and Igλ-expressing cells?**

  Outcome: Selection for and against individual gene segments was observed and the selection profile identified in health was also true for SLE.

- **Is there evidence of differences between health and SLE in receptor editing?**

  Outcome: No selection for or against *IGKJ* gene segments was observed in the productive IGK rearrangements from Igκ- or Igλ-expressing cells in health and SLE. Both health and SLE non-productive *IGK* sequences exhibited a low frequency of *IGKJ4* rearrangement, potentially an imprint of receptor editing.

- **Is there any difference between health and SLE in the expression of *IGK* rearrangements using different gene segments?**

  Outcome: The two most 3' variable gene segments *IGKV4-1* and *IGKV5-2* were significantly more frequently observed in the expressed repertoire of mature naive B cells from SLE patients compared to health.

- **Is there any difference between health and SLE in the components of the CDR3 regions of Igκ light chains that are expressed compared to those that are productive forbidden and not used?**

Outcome of CDR3 length analysis: Shorter CDR3 lengths were observed in the expressed *IGK* repertoire compared to those observed in the productive forbidden repertoire from Igλ-expressing cells in health and SLE.

Outcome of GRAVY index analysis: More negative GRAVY indices were observed in the expressed *IGK* repertoire compared to productive forbidden *IGK* rearrangements in health and SLE. A significantly more negative GRAVY index was observed in the expressed *IGK* repertoire in healthy CDR3s compared to SLE.

Outcome of the theoretical isoelectric point (pI) analysis: Significantly higher theoretical pIs were observed in the expressed *IGK* repertoire compared to the productive forbidden repertoire when CDR3 regions were 10 and 11 amino acids in length in health. This was true for SLE only when CDR3 regions contained 11 residues. Significantly higher theoretical pIs were observed in the expressed *IGK* repertoire with CDR3 regions 9 residues in length in health compared to SLE.

Outcome of aliphatic index analysis: Significantly lower aliphatic indices were observed in the expressed *IGK* repertoire where CDR3 lengths were 8-10 amino acids in length compared to the productive forbidden *IGK* repertoire in health and SLE.

3.3 Discussion

The application of a high-throughput sequencing technique to PCR-amplified Ig rearrangements of human *IGK* light chains expressed by mature naive B cells allowed for investigation into gene segment rearrangement and selection processes during the development of the B cell light chain repertoire in health and SLE. These data suggest that the defect underlying the skewed Ig repertoire previously reported in SLE lies in regulation of light chain expression, rather than in selection and regulation of inherent VJ rearrangement events.

The study of non-productive rearrangements of Ig gene segments allows investigation of an *IGK* repertoire that is not biased by selection because rearrangements cannot be expressed. They reflect the processes that are inherent to the development of a lymphocyte receptor and are governed by several factors; some random and some non-random. The random element of V(D)J recombination involves stochastic processes that diversify the final product of rearrangement through the addition and removal of nucleotides at rearrangement junctions, mediated by TdT and exonuclease activity. As a result of these random events, two thirds of rearrangements are non-productive (out-of-frame) and result in no functional protein. The non-random factors are those that affect the rearrangement of different gene segments differently. V(D)J recombination is initiated by recognition of the recombination signal sequences (RSS) flanking each Ig gene segment by the enzymes catalysing the process; RAG-1 and -2 proteins that introduce double-stranded DNA breaks. RSSs at individual loci of the Ig gene family (*IGH, IGK* and *IGL*) have pre-determined orientations, and conserved composition such that rearrangement can only take place between gene segments with complementary sequences, thereby obeying the 12/23 rule (Bassing et al., 2000). Although they are largely conserved, some RSSs have minor differences that affect recombination frequency (Marie-Paule Lefranc, 2002). Additionally,

the existence of germline-encoded promoters and transcriptional enhancers embedded within antigen receptor loci regulate the accessibility of Ig gene segments and contribute to the rate of Vκ and Jκ gene segment recombination (Bassing et al., 2000; Fulton and Van Ness, 1993; Jung and Alt, 2004).

The high-throughput sequencing data reported in this chapter first compared the *IGKV* rearrangement frequency in non-productive *IGK* rearrangements from the genomic DNA of Igκ- and Igλ-expressing mature naive B cells in both health and SLE. With the exception of infrequently rearranged segments *IGKV3D-20* and *IGKV2-24*, which were seen at higher relative frequency in SLE compared to health, the same inherent biases in *IGKV* rearrangement frequency are apparent in both health and SLE. This may suggest that the skewed usage of *IGKV* rearrangements previously observed in SLE are not the result of a defect in the inherent mechanisms governing VJ recombination in *IGK* light chains.

Forbidden productive sequences were considered to be of interest because they were rearranged productively yet cell fate decisions were made to not proceed with expression of Igκ but to initiate rearrangement at the *IGL* loci. Forbidden sequences are present at 20-25% in health and SLE as shown in this thesis. In mice, such productively rearranged but inactivated *IGK* light chain alleles have been shown to be present in 47% of IgM^+ Igλ-expressing B cells (Retter and Nemazee, 1998). The relatively higher proportion of forbidden sequences in mice is expected since the percentage ratio of Igκ:Igλ light chain expression is 95:5.

Comparison of forbidden rearrangements with those productively rearranged from Igκ-expressing B cells allowed for investigation into the processes governing selection for rearrangements involving specific *IGKV* gene segments. This analysis revealed that

selection both for and against individual segments was apparent. Selection against rearrangements of *IGKV1-37* and *IGKV3-7* is expected. *IGKV1-37* is non-functional due to the replacement of a cysteine residue with glycine at position 104 in framework region 3 (Marie-Paule Lefranc, 2002). *IGKV3-7* has a mutation in an acceptor splice site.

It was not clear what aspect of functionality was responsible for selection against the other segments; this could be based on ability of the new light chain to associate with heavy chain, or the production of an unwanted specificity. Little is known about mechanisms regulating V_H:V_L pairing, mainly due to the fact that any unsuccessful pairings would not be expressed by the cell. Only successful *IGH:IGK/L* pairs would subsequently be exposed to processes regulating antigenic selection at the cell surface since it is the six CDR loop regions contributed by both the heavy and light chains that determine the overall conformation and antigen interface of the binding site (Chailyan et al., 2011). It has been suggested that half of all productively rearranged Ig light chains can pair with IGHμ during development in the small pre-B II cell (Melchers et al., 2000).

The equivalence of this process in health and SLE however demonstrates that the mechanisms involved in identifying that these segments have been rearranged at the pre-B stage and the subsequent cell-fate decision to move to rearrange at the *IGL* locus is the same in SLE and health with respect to segment selection.

In contrast to the comparisons of genomic DNA that showed no differences between health and SLE, biases in the *IGK* repertoire in SLE were apparent by analysis of *IGK* gene expression in the cDNA of the same cells from which genomic DNA had been isolated. This suggests that the mechanisms regulating the expression of *IGK* rearrangements are responsible for the skewed light chain repertoire in reported SLE, and not the process of rearrangement and subsequent segment selection per se. This was not a consequence of

primer bias during PCR because the same set of *IGKV* segment primers was used to derive each set of sequences. These data are also consistent with Yurasov and Wardemann who identified a bias towards *IGKV4-1* expression in single mature naïve B cells in SLE by analysis of gene expression. This study did not include an analysis of genomic DNA and therefore would not have identified that the bias was only at the level of gene expression (Wardemann et al., 2003).

Over-representation of *IGKV4-1* has also been reported following analysis of rearranged *IGK* genes from the genomic DNA from single $CD19^+$ B cells of an SLE patient (Dorner et al., 1998). This study also provided evidence for the clonal expansion and hypermutation of light chain rearrangements involving *IGKV4-1* (Dörner et al., 1999). These data suggested that *IGKV4-1* segments are indeed expressed and cells expressing this light chain may be readily activated. The *IGKV4-1* segment has also been shown to be associated with F4+ B cell lines generated from 6 SLE patient. F4+ refers to the heavy chain idiotype of the antibody which are almost all IgG and bind dsDNA strongly. F4 expression is present predominantly in the serum of SLE patients at high titres (Manheimer-Lory et al., 1997).

The gene segments *IGKV4-1* and *IGKV5-2* in health are both present with lower frequencies in the cDNA of Igκ-expressing cells than expected expression when comparing to their frequencies in the pool of productive rearrangements in the genomic DNA from the same cells. Regulation of *IGK* expression is dependent in part on the activity of the kappa deleting element (KDE), or RS in mice (Nemazee, 2000; Siminovitch et al., 1985; Siminovitch et al., 1987). The KDE is a non-coding, suppressive rearranging element that is located 25 kb downstream of *IGKC* and recombines to either a *IGKV* segment upstream of the unsuccessful rearrangement or to a non-canonical intronic recombination signal sequence (RSS) between the most downstream *IGKJ* segment and the constant region (Klobeck and Zachau, 1986). This rearrangement event is V(D)J recombinase-dependent, and deletes the constant region and associated enhancers thus inactivating the allele

(Takeda et al., 1993). This would imply that the segments that are productively rearranged but not expressed in Igκ-expressing cells have been KDE inactivated. The KDE has been reported to have lower frequency of rearrangement in SLE compared to health. (Panigrahi et al., 2008). It is possible that the KDE fails to inactivate loci that have been selected against in Igκ-expressing cells in SLE, permitting the 'breakthrough' expression of *IGKV5-2* and *IGKV4-2*.

Data in this chapter also identified that *IGKV5-2* is selected against during development at the point where cells make the fate decision to express Igκ protein or to move on and rearrange *IGL*. In contrast, *IGKV4-1* appears not to be selected against at this point, though both segments were selected against at the point of expression. This suggests that multiple mechanisms are involved in identifying and regulating the *IGK* gene expression, presumably including those that initiate and regulate KDE activity. These may act at different times during B cell development. Their activity is essential because most B cells in the blood contain multiple light chain gene rearrangements, and each cell has the ability to produce two different light chains if rearrangements were productive; one encoded by each allele. Restriction to the expression of one light chain is known as isotypic exclusion, though active selection processes that govern this process are elusive. Failure to effectively regulate can lead to aberrant expression, potentially allelic inclusion and a less stringently regulated light chain repertoire.

Comparison of the CDR3 regions associated with productively rearranged *IGK* sequences in the expressed naive *IGK* repertoire with forbidden productive *IGK* rearrangements revealed some interesting differences regarding the components of the antigen binding sites that may influence selection for or against expression. Previous studies have predominantly focused on the CDR3 regions of Ig heavy chains as they form the physical centre of the combining site of the antibody. Longer IgH CDR3s have been associated with DNA-binding antibodies, and shorter CDR3s have been associated with class-switched, memory

B cell populations (Crouzier et al., 1995; Radic and Weigert, 1995; Wu et al., 2010). Changes in length of the CDR3 therefore affects the potential for contacts in the site and overall conformation of the binding domain. Further studies have demonstrated that the IgH CDR3 may distinguish polyspecific from monospecific combining sites in natural and antigen-induced antibodies (Chen et al., 1991).

Variations in these parameters have been described after analysis of the IgH CDR3 region in naive, IgM memory and switched memory B cell subsets. Naive B cells were associated with a downward trend in hydrophobicity and a decreased aliphatic index compared to memory B cell groups. In addition, the *IGH* CDR3 regions of IgM memory B cells exhibited lower aliphatic indices compared to those of switched memory B cells. The analyses of CDR3 characteristics illustrated in this chapter suggested that despite the relatively short length of Ig light chain CDR3 regions in comparison to CDR3 regions of *IGH* chains, subtle but significant differences are associated with the likelihood of a rearranged *IGK* light chain being expressed or not. CDR3s of the expressed *IGK* repertoire tended to be less hydrophilic in SLE compared to health, and CDR3s of the most commonly expressed length (9 amino acids) tended to hold lower theoretical pIs in SLE compared to health. These findings may suggest that the features of the CDR3 in the expressed SLE *IGK* repertoire may contribute to the skewed range of antigens targeted by SLE B cells.

The caveat of the data generated by high-throughput sequencing was that different numbers of sequences were obtained from each individual and for each of three groups of sequences isolated from Igκ-and Igλ-expressing populations of cells in both DNA and cDNA samples. Time and resources limited the study to one high-throughput sequence analysis, though if an experimental repeat was feasible, DNA and RNA would have been isolated from the same number of cells across all samples, rather than the same number of cells within each sample where variability was present (Appendix 1.)

This chapter suggests that the skewed Ig repertoire observed in SLE is the result of the dysregulation of *IGK* gene expression at a genetic level. The following chapter will investigate a role for the KDE in regulating *IGK* gene rearrangements and will consider the consequences of KDE rearrangement in the regulation of Igκ light chain gene expression.

Chapter 4
Analysis of the rearranging activity of the kappa deleting element in populations of CD19^{+} B cells in health and SLE

4.1 Introduction

Following successful heavy chain rearrangement in the developing pro-B cell, the light chain genes generally rearrange in a sequential allelic order, occurring at the *IGK* locus before the *IGL* locus becomes available (Lewis et al., 1982). A consequence of sequential VκJκ rearrangements on one or both alleles at the *IGK* locus, and an additional one or two rearrangements at the *IGL* locus in some B cells, is that a circulating population of B cells may carry multiple Ig light chain rearrangements. These rearrangements are either expressed by the cell after successful pairing with a functional heavy chain, or they are not used due to functional inactivation of the unwanted allele. Alleles of *IGK* that become inactivated include those that have been selected against based on an aspect of their functionality and those not in the incorrect genetic reading frame. Mechanisms governing rearrangements at the light chain loci are therefore important to the maintenance of phenotypically silent light chain rearrangements (isotypic exclusion) and ultimately central B cell tolerance.

The sequential order of light chain gene rearrangements in developing human B cells was first proposed from studies of transformed human B cell lines and leukemic lymphocytes. These early studies largely involved Southern blot analyses of Ig light chain rearrangements and identified that Igκ$^+$ pre-B cell lines or leukemic cells carried alleles of *IGL* in germline configuration, whereas Igλ$^+$ B cells all carried *IGK* alleles that had either undergone rearrangement or had been deleted by an unknown second mechanism governing hierarchical light chain rearrangements (Hieter et al., 1981; Korsmeyer et al., 1981; Lewis et al., 1982). This mechanism was later identified as rearrangement of the kappa deleting element (KDE), and it was demonstrated that 98.5 % of Igλ$^+$ human leukemic B cells carried *IGK* alleles that had undergone Cκ deletion mediated by KDE rearrangement (Beishuizen et al., 1994; Klobeck and Zachau, 1986). These findings were supported by

analysis of VκJκ joints from amplified V gene rearrangements from the Igκ$^+$ and Igλ$^+$ single cell FACS-sorted healthy human B cells (Bräuninger et al., 2001).

The KDE is a recombining element located approximately 25 Kb downstream of the *IGK* constant (C) region and its rearrangement reflects functional inactivation of kappa light chains (Klobeck and Zachau, 1986). KDE rearrangements are considered a normal part of the developmental process in B cells and can take place before rearrangement at the *IGL* loci (Hieter et al., 1981; Korsmeyer et al., 1981). Rearrangement of the KDE takes place to either an isolated, site-specific palindromic RSS (CACAGTG) located within the Jκ-Cκ intron located 917 base pairs 3' to the last Jκ segment, or to an upstream unrearranged Vκ gene segment if available (Graninger et al., 1988; Klobeck and Zachau, 1986; Siminovitch et al., 1985). The former process results in the retention of VκJκ rearranged joints within the cell (Bräuninger et al., 2001), and is accompanied by deletion (or very rarely inversion) of *IGKC* and the intronic (iEκ) and 3' (3'Eκ) enhancers flanking this gene segment. The latter deletes the VκJκ joint. KDE rearrangements to these two sites takes place at frequencies of approximately 75% and 25% respectively (Beishuizen et al., 1994; Sakano et al., 1979).

The KDE has a murine equivalent; the recombination sequence (RS) (Daitch et al., 1992; Durdik et al., 1984). Evidence detailing the mechanism of KDE rearrangements in both humans and mice strongly supports the notion that they are evolutionarily conserved events with an important functional role. Despite the marked differences in the ratio of Igκ: Igλ light chain expression between humans and mice (60:40% and 95:5% respectively), it has been demonstrated that the mouse RS sequence rearranges to an identical CACAGTG 'acceptor site' sequence within the Jκ-Cκ intron, and that albeit relatively small, mouse Igλ-producing B cells frequently carry deletions of Cκ (Durdik et al., 1984).

Thus, KDE rearrangements prevent a series of activities that may take place at *IGK* loci. Firstly, KDE rearrangement prevents the expression of *IGK* gene segments that have been selected against. Secondly, KDE rearrangement prevents further rearrangement events from taking place on the same allele in developing B cells. Thirdly, KDE inactivation of *IGK* loci prevents somatic hypermutation of variable regions should the cell carrying the non-productive rearrangement subsequently become activated because this process is dependent upon the intronic enhancer iEκ. Additional studies demonstrated that both enhancers play a role in activating VκJκ rearrangement events, since double-mutant B cells exhibit no *IGK* rearrangement but can rearrange and express rearrangements of *IGL* (Inlay et al., 2002).

It has previously been suggested that rearrangements of the KDE mirror the level of receptor editing taking place at the light chain loci, and it was proposed that failure of KDE rearrangement, or the equivalent RS element in mice, may be associated with autoimmune disease states (Panigrahi et al., 2008). This group observed lower KDE rearrangement frequency in populations of B cells from patients with T1D and SLE compared to healthy controls. Data were interpreted as decreased levels of receptor editing taking place at the light chain loci in these autoimmune conditions, ultimately suggesting failure in central B cell tolerance.

This chapter describes the use of a quantitative RT-PCR assay to quantify rearrangement events of the KDE on *IGK* alleles of human $CD19^+$ Igκ- and Igλ-expressing B cells to investigate whether the KDE rearrangement process is inefficient in SLE, and whether it therefore may be implicated in the failure to inactivate expression of certain *IGKV* gene segments observed in SLE compared to health, as described in Chapter 3.

4.1.1 Experimental aims

The aims of the analyses described in this chapter were to investigate the nature of the rearrangements of the human KDE in healthy and SLE B cells. The following questions were examined.

- Is there any difference in the relative quantification of KDE rearrangements to the J_κ-C_κ intronic RSS in gDNA isolated from FACS-sorted populations of CD19+ Igκ/Igλ B cells isolated from healthy donors and SLE patients?
- Does the KDE rearrange with equivalent frequency in health and SLE by its alternative rearrangement mechanism – to unrearranged, upstream Vκ gene segments?
- At the molecular level, are there differences in the nature of the KDE rearrangement junctions and breakpoints formed following rearrangement to the J_κ-C_κ intronic RSS between healthy and SLE B cells?

4.2 Results

4.2.1 Development and validation of a multiplex quantitative real-time PCR assay to quantify KDE rearrangements.

A quantitative real-time PCR assay was used to investigate the frequency of KDE rearrangement to the Jκ-Cκ intron (iRSS-KDE) in polyclonal populations of FACS-sorted B cells in health and SLE. This analysis used primers that have previously been applied in a similar quantitative real-time PCR to determine the relative rearrangement activity of the KDE in human subjects (Panigrahi et al., 2008). A schematic outlining the location of the primer binding sites on the *IGK* light chain is depicted in Figure 4.1 A.

The initial experiment used the iRSS-KDE primers to verify that the genomic DNA of a Nalm-6 cell line could be used to efficiently quantify iRSS-KDE rearrangements over a range of concentrations. Nalm-6 gDNA is isolated from a pre-B cell line known to have KDE rearrangements on both *IGK* alleles so represents a positive control for iRSS-KDE quantification. This was performed in order to verify that these primers would be suitable for accurate detection of iRSS-KDE rearrangements from gDNA isolated from varying numbers of cells; the variable relating to the numbers of FACS-sorted B cells obtained for each sample.

Amplification of the iRSS-KDE rearrangement was standardised relative to the expression of the endogenous reference gene RNaseP in a multiplex qRT-PCR reaction. RNaseP is known to be present in two copies in a diploid genome, so can therefore be used for relative quantitation of copy number targets of the iRSS-KDE rearrangement. Having the two primer sets in the same PCR reaction allowed for compensation for technical variability within each PCR reaction, such as pipetting errors. It also compensated for variabilities in starting DNA material since different cell numbers were sorted into each subset; a

reflection of the relative proportions of different subsets within individual $CD19^+$ B cell populations.

The amount of genomic DNA in each sample could be quantified by the relative Ct value obtained for iRSS-KDE rearrangement relative to the Ct values assigned for RNaseP within the same reaction. Serial dilutions of Nalm-6 genomic DNA were made and iRSS-KDE rearrangements were quantified relative to RNaseP. Quantification was always performed in triplicate at each concentration of serially diluted Nalm-6 DNA. Doubling dilutions of Nalm-6 DNA was reflected in one-cycle increments in cycle threshold (Ct) values obtained for iRSS-KDE rearrangement quantification. This was also mirrored by one-cycle increments of RNaseP Ct cycle every dilution (Figure 4.1B).

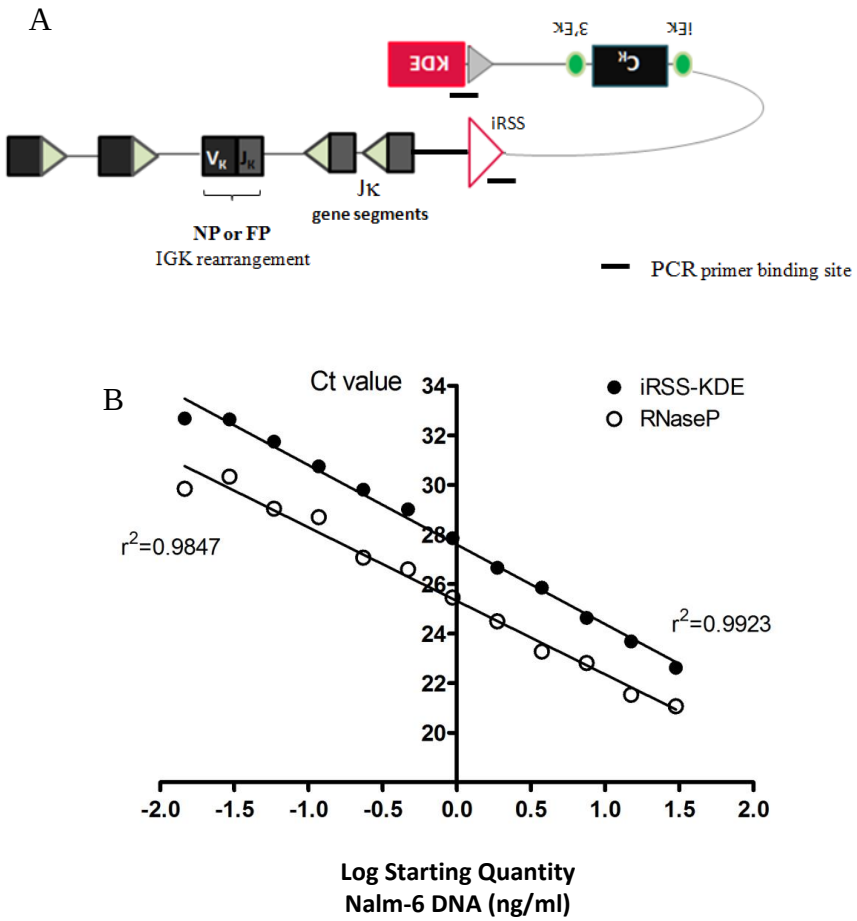


Figure 4.1 Quantification of iRSS-KDE rearrangements. **(A)** Schematic outlining the rearrangement of the KDE to the recombination sequence located in the Jκ-Cκ intron and resulting deletion of Cκ and the enhancers iEκ and 3'Eκ in green. Rearrangement of the KDE takes place if a non-productive rearrangement is produced (NP) or if a productive rearrangement has been selected against and not used; forbidden productive (FP). The black lines represent the approximate location of the primer binding sites for quantification of iRSS-KDE rearrangements following alignment of their RSSs. **(B)** Quantification of iRSS-KDE rearrangements in serially diluted Nalm-6 genomic DNA relative to the housekeeping gene RNaseP in a multiplex quantitative real-time PCR reaction.

Analysis of the real-time PCR data involved comparing the Ct values of the iRSS-KDE rearrangement with RNaseP, and then comparing with the Nalm-6 calibrator by the following calculation.

This method for analysis of RT-PCR data is referred to as the $2^{-[\Delta][\Delta]\,Ct}$ method, where:

ΔΔ Ct = [Δ]Ct of sample normalised to RNaseP – [Δ]Ct of Nalm-6 calibrator normalised to RNaseP

The [Δ][Δ]Ct calculation was a valid analysis for this iRSS-KDE rearrangement quantification as the amplification efficiencies of the target and the endogenous reference RNaseP were approximately equal.

Several quality control parameters were established prior to relative quantification of iRSS-KDE rearrangements of genomic DNA. Previous data from Professor Spencer's lab had determined that Ct scores greater than 35 cycles gave inaccurate quantification of the target gene, so only samples where Ct scores of less than 35 were obtained were considered for accurate iRSS-KDE quantification. Furthermore, where a Ct difference greater than one full cycle was observed among triplicates, the whole data point was excluded. Where a 0.3 cycle difference in Ct score was obtained for two samples within a triplicate, the mean of these two values was taken, and the third score omitted, providing this Ct score was within the 1 cycle permitted for inclusion of the data point.

The quality control data shown in Figure 4.1 confirm that the use of Nalm-6 is reliable as a sample calibrator for iRSS-KDE rearrangement quantification in the genomic DNA isolated from B cells. For the subsequent real-time PCR assays, genomic DNA from Nalm-6 was used at a 1:100 dilution as a data calibrator. At this concentration, all samples could be quantified in the exponential phase of amplification, thereby providing the most precision for iRSS-KDE quantification.

4.2.2 Relative quantification of the rearrangement frequency of the KDE to the Jκ-Cκ intronic RSS in CD19$^+$ B cells of healthy controls and SLE patients.

After optimisation, the previous assay used for quantification of KDE rearrangement to the iRSS was applied to genomic DNA isolated from FACS-sorted viable CD19$^+$ Igκ$^+$ / Igλ$^+$ PBMCs of 10 buffy coats from anonymous healthy controls and 15 female SLE patients. As described in Chapter 3 (section 3.1.2.1), following cell surface immunostaining for Igκ and Igλ light chains, flow cytometry analysis revealed that in some SLE samples, varying proportions of CD19$^+$ cells appeared to harbour both light chains. This observation is investigated in Chapter 5, though for the purpose of this iRSS-KDE quantification by RT-PCR, all cells sorted were gated upon a high expression of Igκ or Igλ and cells falling within intermediate expression levels of either light chains were excluded. An example of this is depicted in Figure 4.2 A.

All SLE patients studied were serum ANA positive, though the population displayed a broad spectrum of clinical manifestations and had had different disease duration (Appendix 2). The relative quantity of iRSS-KDE rearrangement joints will be expressed in terms of the delta delta CT (ΔΔCt) and will be referred to as a relative copy number (RCN), where the quantification of iRSS-KDE is standardised relative to RNaseP in the multiplex reactions, and then calibrated to the ΔCt for iRSS-KDE rearrangement status quantified from the positive Nalm-6 cell genomic DNA control which was always adjusted to '1'. This represents a 1:1 ratio between RNaseP and rearranged KDE since Nalm-6 carries two copies of iRSS-KDE rearrangement per genome. Quantification of iRSS-KDE rearrangements within CD19$^+$ Igκ$^+$ B cells showed no difference between health and SLE (Figure 4.3 B). However, iRSS-KDE rearrangement frequency within populations of

$CD19^{+}Ig\lambda^{+}$ B cells isolated from SLE patients was significantly reduced compared to healthy controls (p = 0.0032).

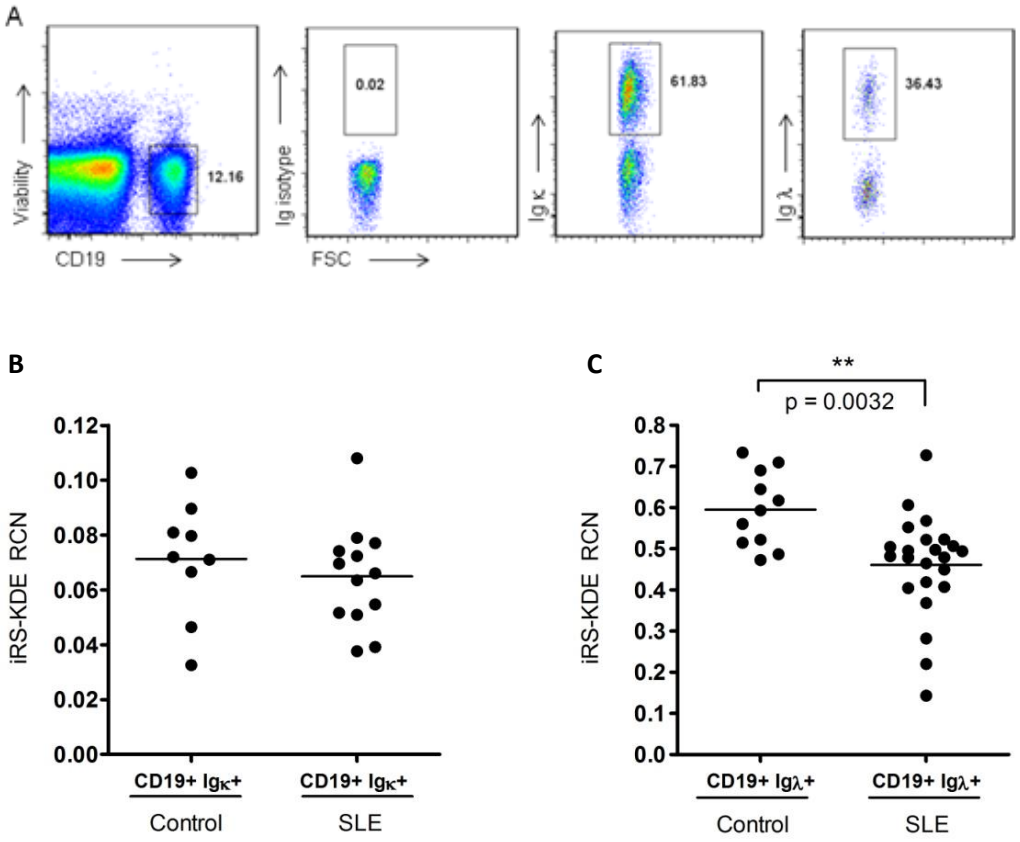


Figure 4.2 Relative quantification of the rearrangement frequency of the KDE in populations of $CD19^{+}$ B cells from healthy controls or SLE patients. **(A)** Outline of the FACS gating strategy used to sort viable $CD19^{+}$ Igκ- and Igλ-expressing B cells. **(B)** Relative quantification of the iRSS-KDE in $CD19^{+}$ $Ig\kappa^{+}$ B cells isolated from healthy controls and SLE patients. **(C)** Relative quantification of the iRSS-KDE in $CD19^{+}$ Igλ+ B cells isolated from healthy controls and SLE patients. Each dot relates to an individual healthy/patient sample. Statistical analyses were performed by Mann-Whitney tests assuming non-equal variance where * p<0.05, **p<0.01, *** P<0.001.

4.2.3 Relative quantification of the rearrangement frequency of the KDE to the Jκ-Cκ intron 5' of the RSS in CD19+ PBMCs of healthy controls and SLE patients.

The significantly lower iRSS-KDE rearrangement frequency observed in SLE $CD19^{+}Ig\lambda^{+}$ B cell populations compared to equivalent $CD19^{+}$ $Ig\lambda^{+}$ B cell populations in health raised questions regarding the nature of the KDE rearrangement. As previously described, the KDE most frequently rearranges to the iRSS, but can rearrange to an unrearranged upstream Vκ gene segment 5' of the iRSS. To ask whether the reduced rearrangement activity of the KDE to the iRSS observed in SLE compared to health was due to more frequent KDE rearrangement to unrearranged Vκ segments, RT-PCR primers were designed to target a portion of the Jκ-Cκ intron, 5' of the iRSS in the same genomic DNA samples from which iRSS-KDE rearrangement frequency had been determined (Figure 4.3 A). If quantification of this portion of the intronic sequence 5' of the iRSS was also less frequent in SLE compared to health, it may suggest that this is the explanation for the reduced frequency of iRSS-KDE rearrangements observed. Quantification of KDE rearrangements to this intronic sequence 5' of the iRSS was equivalent in both $CD19^{+}$ $Ig\kappa^{+}$ and $CD19^{+}$ $Ig\lambda^{+}$ populations of B cells in health and SLE (Figure 4.3 B and C).

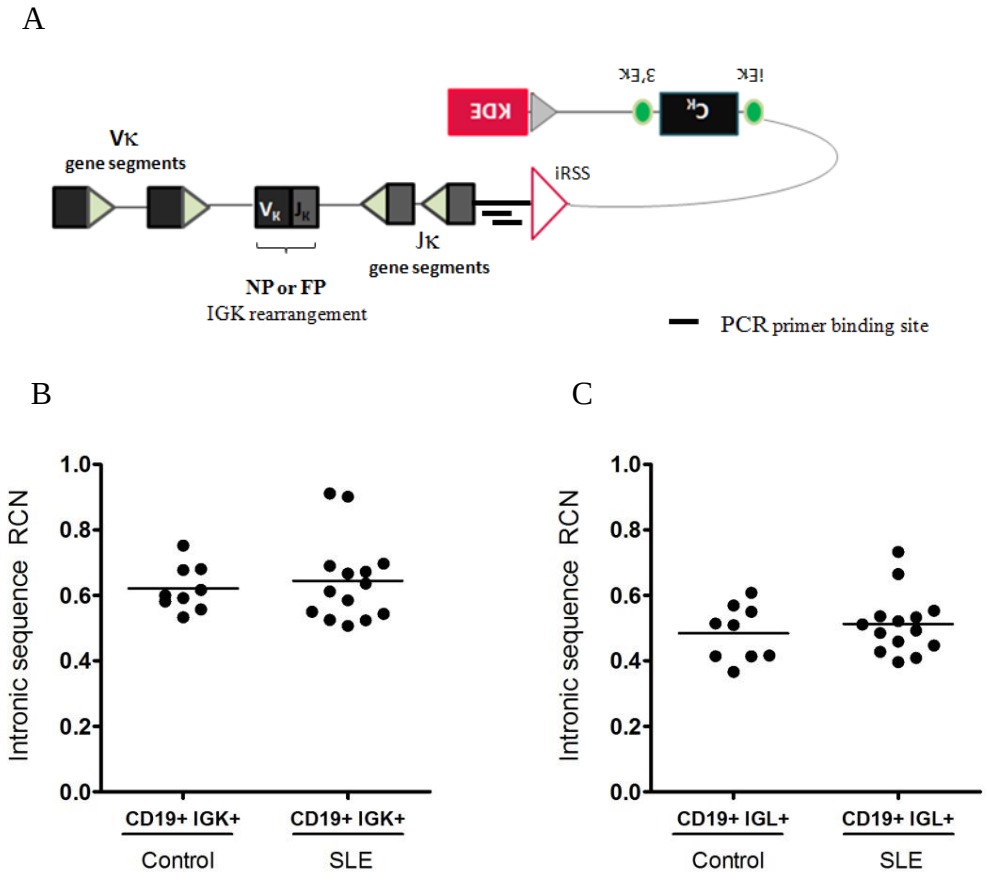


Figure 4.3 Quantitative RT-PCR for relative quantification of the rearrangement of the KDE to an upstream Vκ gene segment. **(A)** Schematic representation of primers used to quantify a portion of the intron 5' of the iRSS. **(B)** KDE rearrangement frequency quantified in FACS-sorted $CD19^+$ $Ig\kappa^+$ B cells in healthy controls and SLE. **(C)** KDE rearrangement frequency quantified in FACS-sorted $CD19^+$ $Ig\lambda^+$ B cells in healthy controls and SLE. Statistical analyses were performed by Mann-Whitney tests assuming non-equal variance.

This section provides evidence to suggest that the reduced frequency of iRSS-KDE rearrangements observed in SLE compared to health was not the result of the KDE rearranging more frequently to an upstream unrearranged Vκ gene segment in SLE.

4.2.4 Sequence analysis of the iRSS-KDE rearrangement junctions from PCR amplified and cloned genomic DNA isolated from CD19$^+$ Igλ^+ PBMCs of healthy controls and SLE patients.

It was considered possible that the reduced frequency of KDE rearrangements to the iRSS observed in SLE might be due to different sequences of the recombination sites involved in the rearrangement itself. Therefore, KDE rearranged junctions were sequenced and the possibility that there may be differences in the local sequence or breakpoint was investigated. Genomic DNA of CD19$^+$ Igλ^+ B cells from 2 healthy controls and 2 SLE patients was PCR amplified using primers designed to amplify the iRSS-KDE rearrangement (Figure 4.4 A). The DNA from Igλ-expressing B cells was studied because these will contain more iRSS-KDE rearranged junctions than those expressing Igκ. PCR products were then cloned and sequenced. A total of 49 sequences from healthy controls and 63 sequences from SLE patients were obtained and analysed using Gene Jockey software. Each sequence was aligned with the germline Jκ-Cκ intron and germline KDE sequence to identify the breakpoint of rearrangement of the KDE to the intronic RSS. The final nucleotide at the junctional breakpoint was obtained for both the intron and KDE, and the percentage frequency of sequences that expressed this breakpoint was calculated. Examples are given in Figure 4.4 B. Similar trends were observed between health and SLE at both the intronic RSS and KDE breakpoint, with a highly diverse set of breakpoints at both the iRSS and the KDE (Figure 4.5).

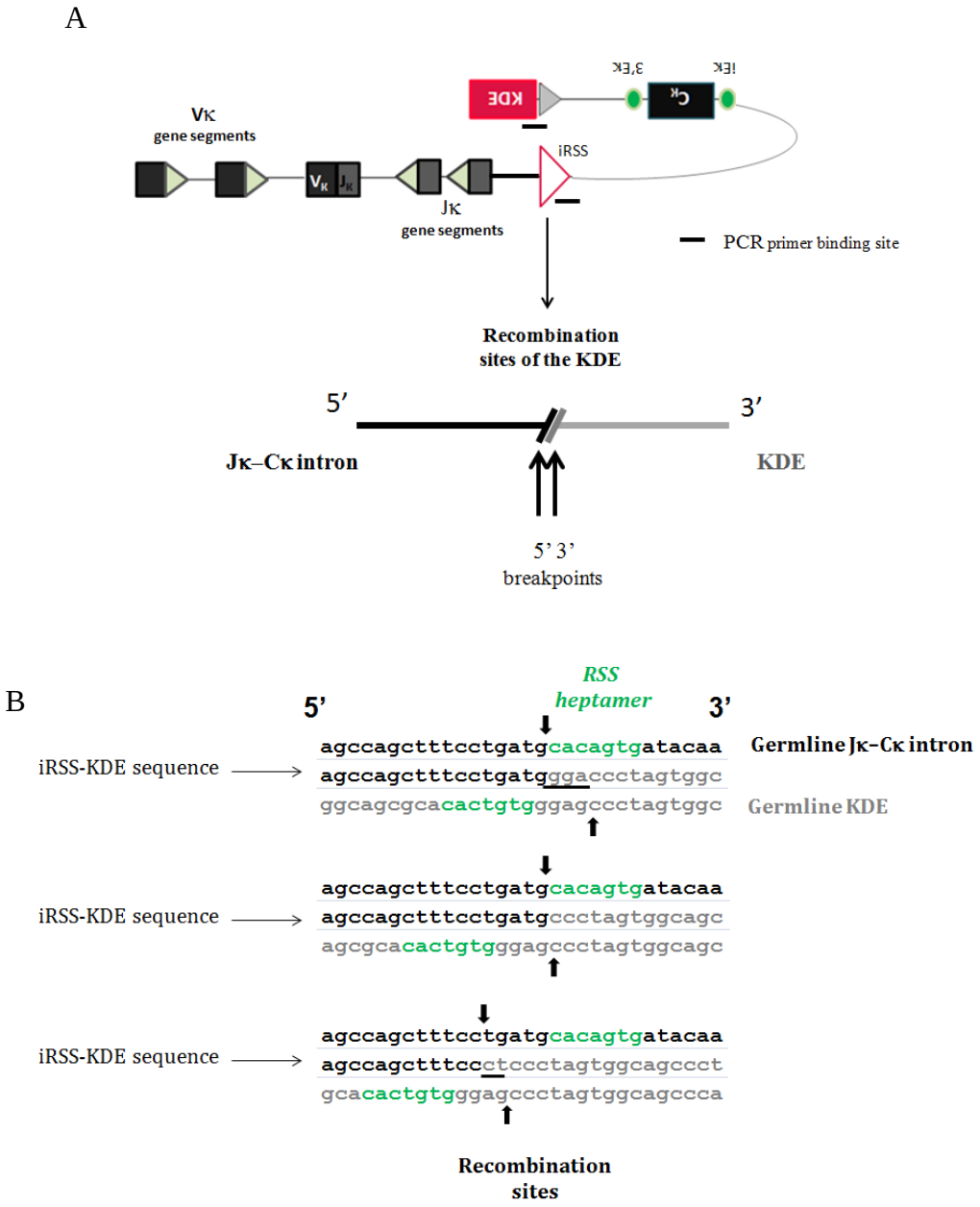


Figure 4.4 Sequence analysis of iRSS-KDE rearrangements. **(A)** Schematic to outline the experimental design. PCR amplification of iRSS-KDE rearrangements was performed using gDNA isolated from FACS sorted CD19+ B cells from two healthy controls and two SLE patients. Following PCR amplification and confirmation of the correct band size by gel electrophoresis PCR products were purified before cloning and subsequent sequencing (Beckman Coulter Genomics). Sequences obtained from these clones were analysed for junctional variability within the iRSS-KDE rearrangements amplified. **(B)** Representative example sequences of a rearranged iRSS-KDE junctions aligned with germline Jκ-Cκ intron and KDE sequences. The black arrows represent where the germline Jκ-Cκ intron ends and the germline KDE sequence begins. In this way the frequency of sequences terminating at specific nucleotides and any addition of N-nucleotides could be assessed.

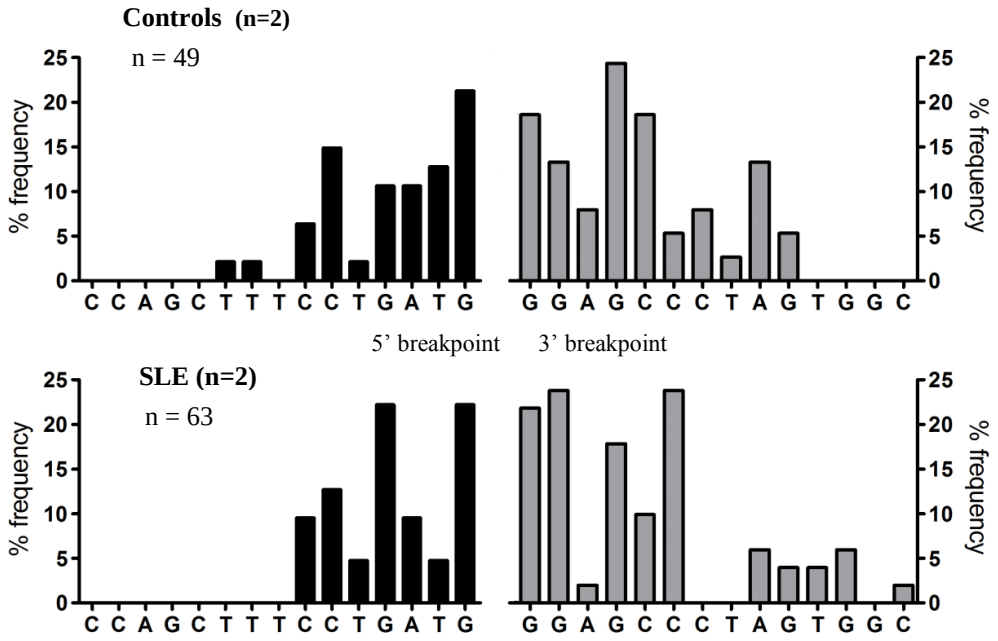


Figure 4.5 iRSS-KDE rearrangement breakpoint analysis. The percentage frequency of sequences that terminate at specific breakpoints after iRSS-KDE rearrangement. PCR products were cloned and sequenced from PCR-amplified genomic DNA isolated from $CD19^+$ $Ig\lambda^+$ populations of B cells in healthy controls (n=2) and SLE (n=2).

Additional information regarding the nature and timing of the KDE rearrangement could be studied through analysis of N-nucleotide addition by TdT at the rearranged iRSS-KDE junction in the pooled control and SLE sequences. TdT varies in its activity during B cell development. Where the addition of P- and N- nucleotides were observed, total additions were pooled because in approximately 25% of cases, an N-nucleotide addition would be indistinguishable from a P-nucleotide. The trend in N-nucleotide addition at the junctions was similar in health and SLE, with approximately 35% of rearranged iRSS-KDE junctions having no N-nucleotide addition (Figure 4.6).

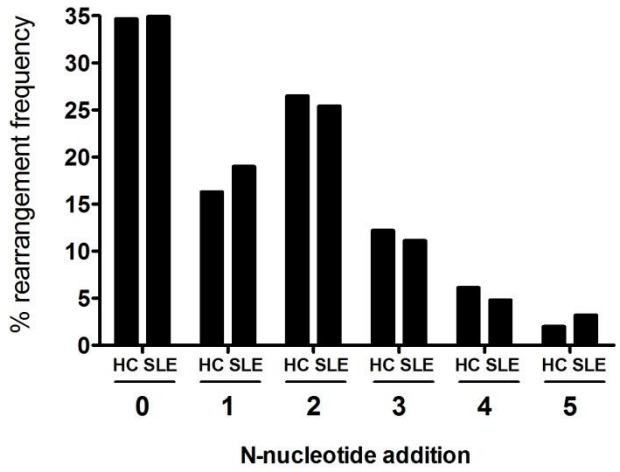


Figure 4.6 N-nucleotide addition at the KDE-iRSS rearranged junction. The percentage frequency iRSS-KDE rearranged sequences from CD19^{+} Igλ^{+} carrying varying N-nucleotide addition in health (HC n=2) and SLE (SLE n=2). Approximately thirty five percent of iRSS-KDE junctions studied lack N-nucleotide addition.

4.2.5 Summary of experimental findings

- **Is there any difference in the relative quantification of KDE rearrangements to the J_{κ}-C_{κ} intronic RSS in gDNA isolated from FACS-sorted populations of CD19+ Igκ/Igλ B cells isolated from healthy donors and SLE patients?**

  Outcome: Relative quantification of iRSS-KDE rearrangements demonstrated that the frequency of KDE rearrangement events to the iRSS is lower in CD19^{+}Igλ^{+} populations of B cells in SLE compared to healthy controls.

- **Does the KDE rearrange with equivalent frequency in health and SLE by its alternative rearrangement mechanism – to unrearranged, upstream Vκ gene segments?**

  Outcome: Relative frequency of a conserved intronic sequence located in the Jκ-Cκ intron but 5' of the iRSS is similar in health and SLE, suggesting that the reduced frequency of iRSS-KDE rearrangements in SLE is not due to rearrangement of the KDE to an unrearranged upstream Vκ segment.

- **At the molecular level, are there any differences in the nature of the KDE rearrangement junctions and breakpoints formed following rearrangement to the J_κ-C_κ intronic RSS between healthy and SLE B cells?**

  Outcome: There was no difference in the features of the rearranged junction or local sequence at iRSS-KDE breakpoints in SLE compared to health.

4.3 Discussion

Rearrangements of the KDE serve to silence newly rearranged VκJκ joints that are in the incorrect genetic reading frame or have been selected against, and are thought to accompany ongoing rearrangement events associated with Ig light chains. As previously mentioned, Panigrahi *et al.*, 2008 described the quantification of KDE rearrangement events to its most common site of rearrangement, the iRSS, in populations of B cells from mice and humans, observing lower KDE rearrangement frequencies in B cells from patients with SLE and T1D. They suggested that KDE rearrangements could serve as a biomarker for receptor editing at the light chain loci, and that their findings were reflective of decreased receptor editing by light chains in autoimmunity (Panigrahi et al., 2008).

The first set of data presented in this chapter aimed to verify the observations of Panigrahi *et al.*, 2008 through use of a similar quantitative real-time PCR assay to investigate the rearranging activity of the KDE to the iRSS in genomic DNA samples. The genomic DNA from populations of Igλ-expressing B cells in SLE carried a significantly lower frequency of rearrangements of the KDE to the iRSS, confirming the previous observations (Panigrahi et al., 2008). In contrast to the data of this study however, no difference was observed in the frequency of iRSS-KDE rearrangements quantified in the genomic DNA from $CD19^+$ Igκ-expressing B cells between health and SLE. This may be due to the relatively small data set acquired compared to the previous study where double the number of SLE patients were studied. In addition, fewer iRSS-KDE rearrangements would be expected from Igκ-expressing cells than Igλ-producers, so perhaps a comparatively larger sample size was required to detect subtleties in iRSS-KDE rearrangement frequency in these $Igκ^+$ cells in order to subsequently observe similar trends as those seen by Panigrahi and colleagues.

These data presented in this chapter may provide an alternative interpretation of the lower KDE rearrangement frequency by SLE B cells to that of Panigrahi *et al.* 2008, who suggest

this reflects defective receptor editing in SLE. In addition, the high-throughput sequencing analysis of *IGKV* and *IGKJ* gene segment rearrangements did not identify any evidence of a defect in receptor editing in SLE. This is consistent with a study by Dorner and colleagues, who obtained single peripheral $CD19^{+}$ B cells from a patient with untreated SLE and studied both *IGKV* and *IGKJ* gene segment usage in the productive and non-productively rearranged *IGK* repertoire. They suggested that the increased usage of *IGKV5* compared to a control donor, alongside its over-representation with groups of productively rearranged sequences compared with the non-productively rearranged sequences in the SLE patients, was reflective of intact and more apparent levels of receptor editing compared to health (Dorner et al., 1998).

The potential of the KDE to rearrange to Vκ-associated RSSs other than the iRSS meant that it was important to question whether the lower frequency of iRSS-KDE rearrangement events observed in SLE was due to an increase in KDE rearrangement events to previously unrearranged Vκ gene segments. Additional studies have involved restriction map analyses of *IGKV, J* and *C* gene segments from different B cell lines including Nalm-6, pre-B cells from acute lymphoblastic leukemia cell lines and SU-DHL-6 cells from an IGHμ, Igλ-producing histiocytic cell line (Siminovitch et al., 1985). These experiments showed that 100% (20/20 cells studied) of Cκ segments were lost from Igλ-producing B cell precursors and that 5 of these cases exhibited deletion of all Jκ segments, representing instances where the KDE rearranged to an upstream Vκ segment (25% of KDE rearrangements) (Sakano et al., 1979). Quantification of a conserved portion of the Jκ-Cκ intron, 5' of the iRSS, in both health and SLE aimed to question whether the KDE may have a tendency to rearrange more frequently to an upstream, unrearranged Vκ gene segment in SLE. The equivalence in quantification of this portion of the intron suggests that this is not the case.

The cloning and sequening of iRSS-KDE rearranged junctions from the genomic DNA of Igλ-expressing B cells from 2 SLE patients and 2 healthy controls allowed for comparison of the diversity in KDE rearrangement junctions, and to question the possibility that there may be differences in the local sequence or breakpoint between health and SLE, which may account for the lower frequency of KDE rearrangements observed. The composition of junctions generated by gene rearrangement is related to the activity of TdT that alters in expression during B cell development. TdT-deficient lymphocytes in mice bear no N nucleotides in their *IGH* variable region genes (Komori et al., 1993), and expression of this enzyme is largely restricted to early B cell differentiation in the pro-B cell stages when *IGH* rearrangement predominates (Li et al., 1993). This provides reason for the limited addition of N-nucleotides observed in VκJκ junctions analysed here, since light chain rearrangement only begins in the pre-B cell stage where only residual TdT would be present. The similarity in the DNA sequences involved in the iRSS-KDE rearrangement in SLE and health implies that they are acquired when relevant enzymatic activity is equivalent, and therefore at an equivalent stage in development.

Although the inherent mechanism of a KDE rearrangement is fairly well documented, the signalling events that initiate the regulatory activity of the KDE remain elusive, largely due to the poor understanding of the signalling pathways involved in initiation of light chain rearrangements. One study described the requirement for transcription factors IRF-4 and IRF-8 in orchestrating the initiation of light chain rearrangements in pre-B cells, as the B cells of IRF4,8$^{-/-}$ mice arrest in the cycling pre-B cell stage (Lu et al., 2003). Five years later, it was suggested that light chain rearrangements are driven synergistically by convergent signalling pathways involving the transcription factor IRF4 and signalling through the IL-7R in the developing B cell, though debate lies in the requirement for acquired pre-BCR signalling or attenuated IL-7R signalling (Johnson et al., 2008).

The fact that the conserved CACAGTG heptamer sequence associated with the KDE is identical to the heptamer sequences that flank germline Vκ and Jκ segments (Sakano et al., 1979) suggests that the recombinatorial enzymes, enhancers and promoters that mediate initial rearrangement activity at the *IGK* loci are also responsible for its functional inactivation by the KDE. It would therefore be of interest to study these enzymes, predominantly the RAG gene products and the signalling pathways upstream of them to investigate potential signalling pathways involved in KDE rearrangement initiation.

The data presented in this chapter, alongside the conclusions drawn in Chapter 3 suggest that rearrangement of the KDE to functionally inactivate *IGK* alleles is inefficient in SLE compared to health, potentially allowing for the emergence of an SLE B cell repertoire expressing *IGK* light chains that are typically selected against in health. The following chapter will explore earlier observations involving the potential expression of both Igκ and Igλ light chains by some B cells in SLE (described in section 4.2.2) and consider the more global consequences of a failure in KDE rearrangement frequency.

Chapter 5
Investigating immunoglobulin cell surface light chain expression by $CD19^+$ B cells in health and SLE

5.1 Introduction

Studies in Chapters 3 and 4 of this thesis required identification and isolation of B cells in SLE expressing Igκ and Igλ light chains as published previously by others (Panigrahi et al., 2008). However, as stated in Chapter 3 (section 3.1.2.1) and Chapter 4 (section 4.2.2), this was frequently more complex than initially considered. In this chapter, potential dual light chain expression by SLE $CD19^+$ B cells is investigated.

The monospecificity of B cells, mediated genetically by allelic exclusion of immunoglobulin heavy and light chain genes during V(D)J recombination, is regarded as one of the most evolutionary conserved features of the mammalian adaptive immune system. Allelic exclusion ensures that the expression of productively rearranged immunoglobulin alleles by a new B cell is limited to just one per cell (Vettermann and Schlissel, 2010), and historically is referred to as the 'one B cell – one antibody' paradigm (Nossal, 1958).

Allelic and isotypic exclusion is established at the light chain loci through sequential rearrangement of alleles at the *IGK* loci which must be functionally inactivated by rearrangements of the KDE before the *IGL* loci undergoes rearrangement. This results in the expression of just one light chain isotype, Igκ or Igλ, and just one rearrangement at these loci if both rearrange productively. Isotypic exclusion limits the proportion of B cells expressing both Igκ and Igλ light chains to approximately 3% of B cells in fetal BM and spleen (Pauza et al., 1993). Interestingly, ordered Ig gene rearrangements are not *per se* required for allelic exclusion. Several studies have demonstrated that chickens, sharks and skates all display allelic exclusion despite rearranging their Ig genes simultaneously at various developmental stages (Hsu, 2009; Ratcliffe, 2006).

Exceptions to the ordered rearrangement of Ig light genes where rearrangement of alleles of *IGL* precede those of the *IGK* loci have also been reported. For example, mouse B cell lines have been shown to produce Igλ light chain protein in the absence of *IGK* rearrangement (Berg et al., 1990; Felsher et al., 1991). In addition, the occasional rearrangement of *IGL* genes despite the maintenance of *IGK* alleles in germline states have been described in three human leukemic B cells in a total of seventeen studied (Tang et al., 1991) and EBV-transformed human pre-B cells undergoing differentiation were shown to produce Igλ in the absence of rearranged *IGK* alleles (Pauza et al., 1993).

Several genetic models have been proposed that serve to explain the existence and maintenance of allelic exclusion of Ig genes. In the asynchronous recombination model, the mechanisms that mediate chromatin accessibility and access of RAG proteins to RSSs are proposed to result in the inability to simultaneously recombine two Ig alleles in one cell (Liang et al., 2004). This model in particular is thought to be relevant to the maintenance of isotypic exclusion exhibited by the Ig light chains, where differences in the activation status, RSS recombination efficiencies and the roles of the enhancers and promoters embedded at the *IGK* and *IGL* loci may limit synchronous recombination and ultimately light chain inclusion (Li et al., 2002b; Ramsden and Wu, 1991). Methylation status is also known to contribute to allelic exclusion at the *IGK* loci, since non-rearranged alleles remain methylated and those undergoing recombination are under-methylated (Mostoslavsky et al., 1998).

The stochastic model proposes that allelic exclusion is the direct result of the low probability of generating a productive Ig gene rearrangement that additionally is a functionally pairing Ig chain. The resultant Igκ:Igλ cell surface expression ratio is therefore the result of the different recombination frequencies of *IGK* and *IGL* loci (Coleclough et al., 1981; Ramsden and Wu, 1991). In this way, allelic inclusion is minimised.

The feedback inhibition model proposed that intermediate products produced during V(D)J recombination serve to inherently inhibit the recombination process itself, where signalling through both the pre-BCR and BCR suppress further recombination (Alt et al., 1984). This model also supports the suggestion that second allele rearrangements take place following the generation of a non-productive or non-pairing heavy or light chain gene, as the lack of feedback inhibition signals allows for the continuation of recombination.

The tendency of mammalian B cells to express just one rearranged light chain (Igκ or Igλ) despite having the genetic potential to express more than one simultaneously is dependent upon the mechanisms that preserve isotypic exclusion and prevent the expression of rearrangements that have been selected against. Defective rearrangement of the KDE and the subsequent failure to functionally inactivate rearranged alleles of *IGK* that have been selected against defies the laws of allelic exclusion and therefore may have dramatic consequences for the light chain repertoire of B cells. This chapter explores the features of the light chain repertoire in health and SLE to understand the features of light chain expression observed in Chapters 3 and 4 more clearly.

5.1.1 Experimental aims

The aims of the analyses described in this chapter were to investigate the nature of cell surface immunoglobulin light chain expression by SLE B cells following the observations that several SLE patients appear to express both Igκ and Igλ light chains on their cell surface. These investigations were addressed by asking the following questions:

- Is a tendency to express both Igκ and Igλ light chains by some B cells a rare or frequent observation in SLE?
- Is this apparent dual Igκ and Igλ light chain expression a genuine feature of SLE $CD19^+$ B cells, or is cytophilic binding of Igs, or Ig-ICs bound to the surface of SLE B cells responsible for interfering with immunostaining prior to flow cytometry analysis?
- Is dual light chain cell surface expression stable?
- Could VpreB surrogate light chain expression account for the apparent dual expression of light chains?
- Can rearranged transcripts of *IGK* and *IGL* be detected in the cDNA isolated from single cells of a dual $Ig\kappa^+Ig\lambda^+$ SLE patient?
- Is dual light chain expression a feature of one specific B cell subset?
- Is there a correlation between dual light chain expression and KDE rearrangement activity in SLE?
- Are serum free light chains associated with dual light chain expression by SLE B cells?

5.2 Results

5.2.1 Investigation of cell surface light chain expression by healthy and SLE CD19$^+$ B cells.

Analysis of rearrangement of iRSS-KDE rearrangements in Igκ- and Igλ-expressing B cells described in Chapters 3 and 4 involved isolating genomic DNA from Igκ- and Igλ-expressing B cells. It became apparent during the FACS sorting for these experiments that the profiles of cell surface light chain expression exhibited by some SLE PBMCs was highly variable in appearance, and in many instances, expression of solely Igκ or Igλ light chains by populations of CD19$^+$ B cells could not be identified in. This was overcome by avoiding cases that had unusual Igκ:Igλ-expressing cell profiles (Chapter 3) or by isolating Igκ and Igλ-expressing cells independently of each other (Figure 4.2 A).

To examine this observation more closely, PMBCs were isolated from the whole blood of 58 SLE patients and 26 anonymous healthy controls and were then immunostained for the cell surface expression of CD19, Igκ and Igλ light chains prior to flow cytometry analysis. The variable nature of cell surface light chain expression by SLE B cells meant that strict gating rules were required in order to maintain a consistent determination of the degree of dual expression observed in each control and SLE sample analysed. Firstly, isotype control stains were performed for each sample which allowed for negative quadrant gates to be set according to these. The same quadrant was then applied to the equivalent stained sample, and adjusted to within 0.5 logs if necessary to control for background fluorescence and the potential for non-specific staining. If gating adjustments within this range were required, this was done through determination of the populations of CD19$^+$ as observed in the contour setting, which could indicate the number of CD19$^+$ populations present with respect to light chain expression. Furthermore, to exclude the possibility that SLE cases may differ in their

individual properties including for example rare background fluorescence, all SLE samples stained were accompanied by isotype controls and unstained cells as well as fluorescence minus one (FMO) controls.

With the exception of one healthy individual, 95% of total $CD19^+$ B cells from healthy controls expressed either an Igκ or Igλ light chain; a representative example of the light chain cell surface expression exhibited by a typical healthy control is shown in Figure 5.1 A. In contrast, some SLE samples appeared to have three distinct populations of $CD19^+$ B cells expressing varying degrees of Igκ and Igλ light chains; 58.15% of $CD19^+$ B cells in SLE sample 1 and 19.03% in SLE sample 4 expressed both Igκ and Igλ on their cell surface. Other SLE samples appeared to have two fairly distinct populations of $CD19^+$ B cells expressing either an Igκ or Igλ light chain, though cells within these populations seemed to drift into quadrants where cells within them seem express both light chain of their cell surface. This is illustrated by SLE example 3, where $Ig\kappa^+$ B cell cells seem to drift into the range of Igλ positivity. Some rare SLE patients appeared to express both light chains on almost all of their $CD19^+$ B cell population, as illustrated by SLE example 2 (Figure 5.1 B).

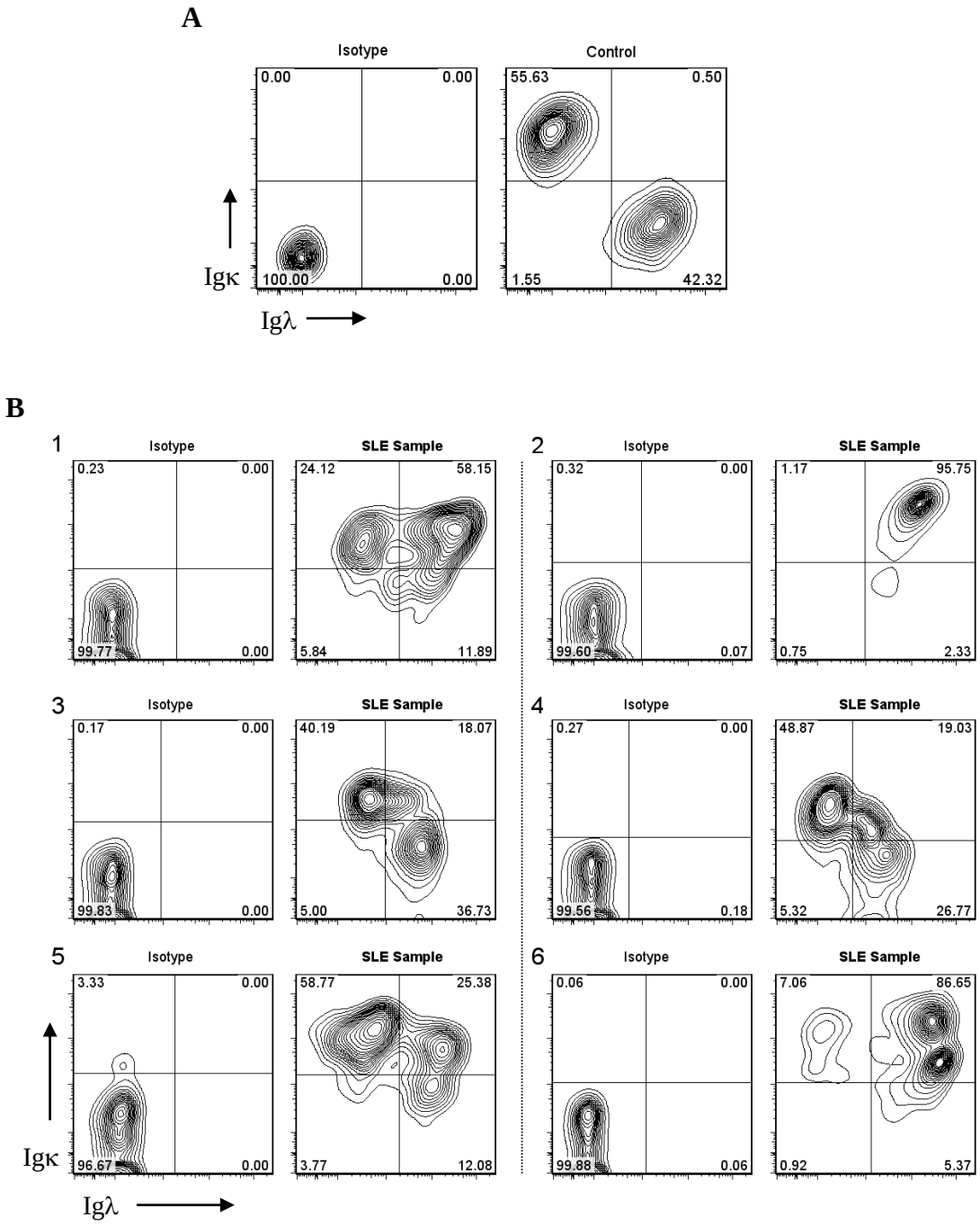


Figure 5.1 Flow cytometry analysis of the cell surface expression of Igκ and Igλ light chains by $CD19^+$ B cells compared to their isotype immunostained controls. **(A)** Typical light chain expression profile by a healthy donor with a ratio of approximately 60:40 Igκ to Igλ. **(B)** Both Igκ and Igλ cell surface expression can be observed in SLE samples 1-6, though to varying degrees.

Comparison of light chain expression in all SLE samples with healthy controls stained by the same method revealed that a significantly higher proportion of $CD19^+$ B cells in SLE exhibited a degree of dual light chain expression compared to the same B cell population in health, suggesting that this phenomenon may be associated with SLE (Figure 5.2). It was next addressed whether any clinical features of disease were associated with this apparent expression of both Igκ and Igλ light chains (clinical information on all SLE patients investigated in this study can be found in Appendix 2). Interestingly, a history of nephritis, including both active and non-active nephritis at the time of blood sample collection, was associated with a significantly higher expression of both Igκ and Igλ light chain by SLE $CD19^+$ B cells compared to healthy control B cells ($p<0.0001$***). SLE patients with no history of nephritis however still harboured significantly higher frequencies of $CD19^+$ B cells expressing both Igκ and Igλ light chains compared to healthy controls ($p<0.0084$**). $CD19^+$ B cells from a disease control patient cohort with granulomatosis with polyangitis (GPA) were similarly stained for cell surface expression of both light chains and were found to express similar levels of Igκ and Igλ as healthy controls (Figure 5.2).

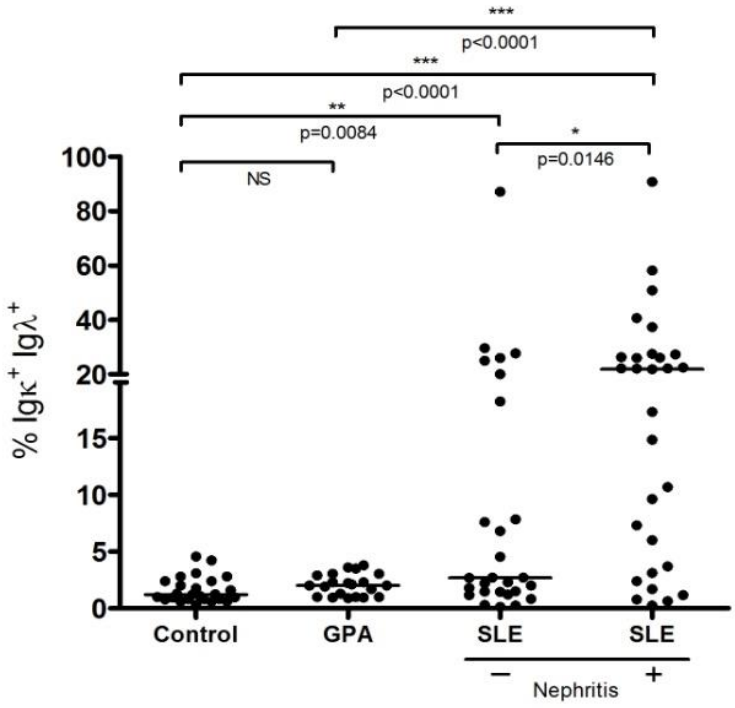


Figure 5.2 Flow cytometry analysis of the proportion of $CD19^+$ B cells expressing both Igκ and Igλ light chains in healthy controls (n=26), disease control GPA patients (n=21) and SLE patients with (n=29) or without (n=27) a history of lupus nephritis (total SLE patients n=56.) SLE patients with a history of nephritis (active or non-active at the time of blood collection) displayed significantly higher populations of $CD19^+$ B cell expressing both Igκ and Igλ light chains on their cell surface compared to non-nephritis SLE patients, and both GPA and healthy control samples.

It was also asked whether the presence of autoantibodies measured in SLE patient serum associated with the likelihood dual light chain expression by $CD19^+$ B cells. The clinical information was retrieved from Lee Meng Choong at St Thomas' Hospital, London, who was responsible for the collection of blood samples and storage of additional clinical information about the patients used in this study. The autoantibodies assessed included the following; anti-dsDNA, anti-ENA, anti-Ro, anti-La, anti-Sm, anti-RNP and anti-aPL. No differences were observed in the frequency of $CD19^+$ B cells expressing both Igκ and Igλ light chains; patients exhibited dual light chain expression irrespective of their positivity for autoantibodies (Figure 5.3).

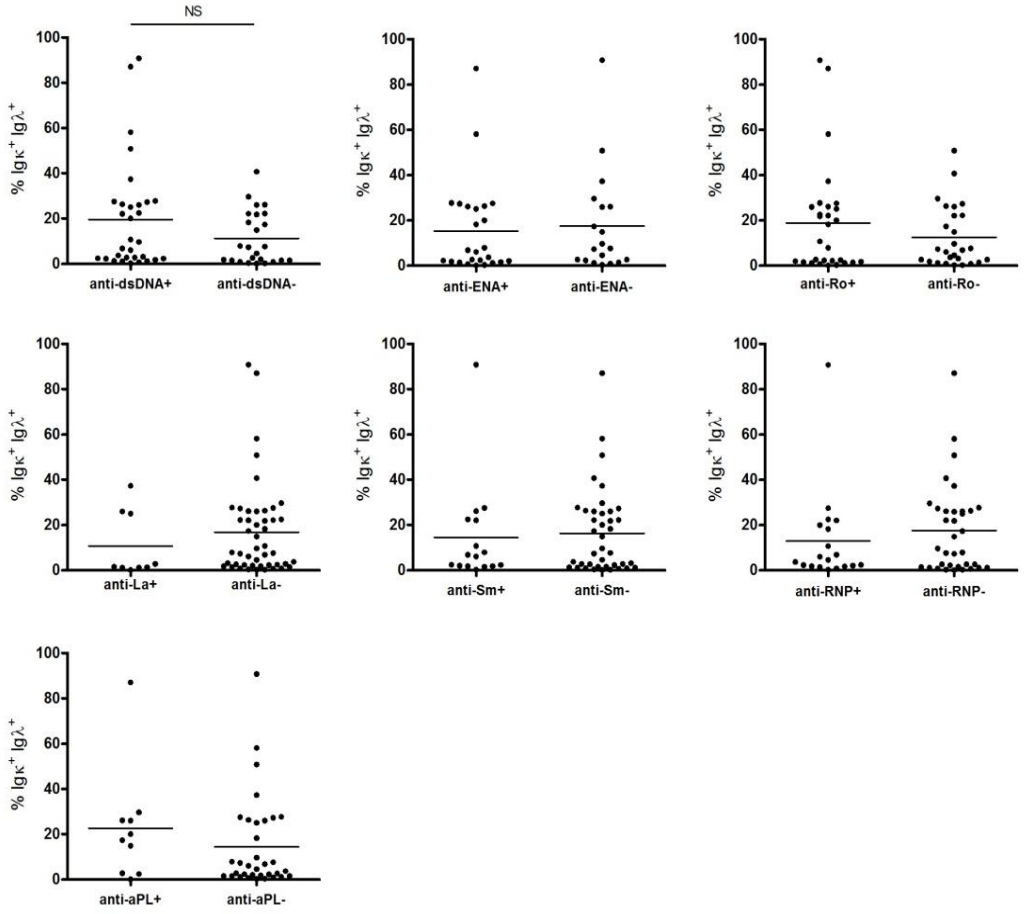


Figure 5.3 Autoantibody titres measured from SLE patient serum and the frequency of SLE CD19$^+$ B cells expressing both Igκ and Igλ light chains. There was no association between whether SLE patients harbour CD19$^+$ B cells expressing both light chains and the presence of autoantibodies.

5.2.2 Investigating the origin of dual light chain expression by SLE CD19$^+$ B cells by flow cytometry analyses.

The observation that B cells appeared to express both light chains to varying degrees in approximately half of all SLE patients studied meant that several areas had to be explored to determine if this observation was a genuine feature of SLE or an artefact. It was first considered, since the hallmark of SLE is the propagation of an autoimmune

immunoglobulin repertoire characterised by the presence of antibodies targeting dsDNA, that SLE B cells may be coated with immunoglobulin-dsDNA immune complexes. This could account for the presence of Igκ and Igλ light chains on the same cell. It is also believed that a contributing factor to the pathogenesis of SLE is the aberrant clearance of apoptotic cells and exposure to self-DNA and histones, leading to the subsequent immune response to autoantigen. Furthermore, lower serum levels of DNase I activity have been observed in active SLE patients compared to healthy donors (Chitrabamrung et al., 1981).

DNase I treatment of PBMCs would remove any bound Ig/dsDNA complexes and sticky DNA molecules that may be present in the SLE samples and ensure they do not interfere with the staining of target cells for surface immunoglobulins. To test this, PBMCs from two SLE patients exhibiting highly distorted light chain flow cytometry profiles were washed prior to incubation with DNase I at a concentration of 0.1 mg/ml at room temperature for 30 minutes, according to manufacturer's instructions. Cells were then washed and immunostained by the typical protocol. DNase I treatment did not alter the distorted light chain expression by the two SLE cases, where after DNase I incubation and subsequent immunostaining, similar proportions of $CD19^+$ B cells expressed both Igκ and Igλ on the cell surface (95.53% and 53.87%) (Figure 5.4).

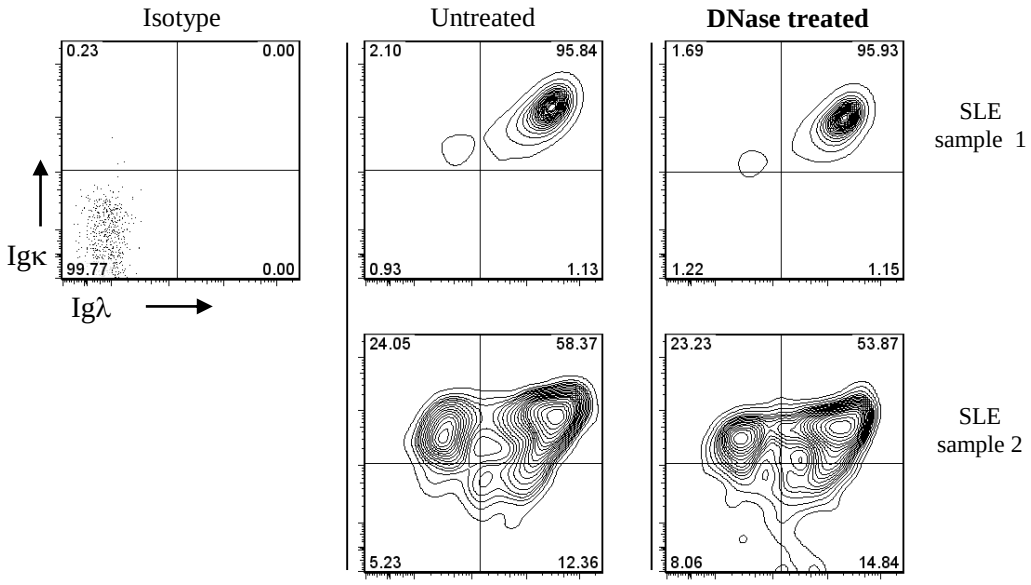


Figure 5.4 Flow cytometry analysis of light chain expression by SLE B cells following incubation with DNase I (0.1 mg/ml) for 30 minutes at room temperature. The large proportions of dual light chain expressing B cell populations observed prior to DNase treatment, remained following DNase incubation.

In order to eliminate the possibility that the detection of both light chains on single cells was a consequence of cytophilic binding or adsorption of serum immunoglobulin to the B cell surface, acidic wash buffers were used to treat the cells prior to immunostaining and flow cytometry analysis. Acid washed have previously been used to detach antibodies from antigen (Kameyama et al., 2007). PBMCs from the same SLE samples as in Figure 5.4 were thawed and washed in 2% FCS. Cells were subsequently washed twice for 5 minutes in either 0.2M glycine 0.15M NaCl (pH 3.0) or RPMI-1640 media (pH 2.5) buffers. Acid washes were followed by two standard wash steps of 2% FCS in PBS. High frequencies of $CD19^{+}$ B cells expressing both light chains were still observed following treatment (Figure 5.5).

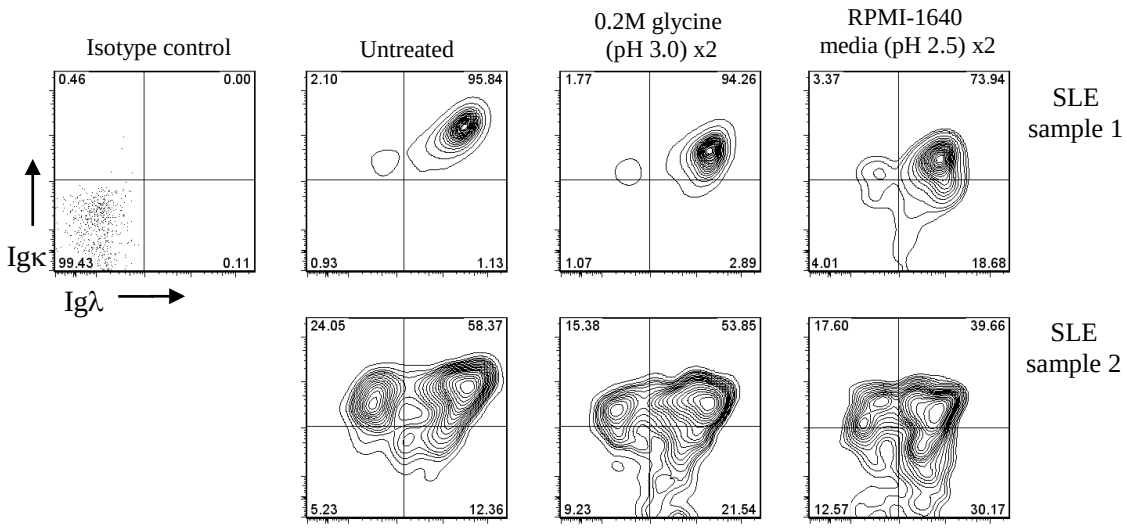
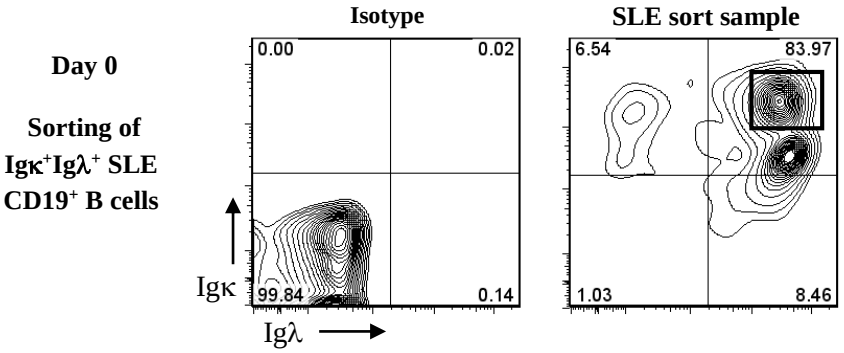


Figure 5.5 Flow cytometry analysis of light chain expression by SLE $CD19^+$ B cells following acid washing in a 0.2M glycine (pH3.0) and an RPMI-1640 media wash (pH2.5) for 5 minutes prior to immunostaining. Large dual-light chain expressing populations of cells were observed following acid washing two different buffers.

To determine whether the expression of both light chains was a stable feature of SLE $CD19^+$ B cells, dual-light chain producing $CD19^+$ B cells from an SLE patient were FACS sorted and co-cultured with allogeneic $CD3^+CD4^+$ T cells isolated from a multidonor buffy coat obtained from the National Blood Service. Cells were stimulated with 50 ng/ml PMA for two days and cells were harvested and re-stained for the expression of CD19, Igκ and Igλ. Cells remained positive for the expression of both light chains after two days of stimulation (Figure 5.6).



Day 2 following co-culture with $CD3^+CD4^+$ allogeneic T cells + 50ng/ml PMA

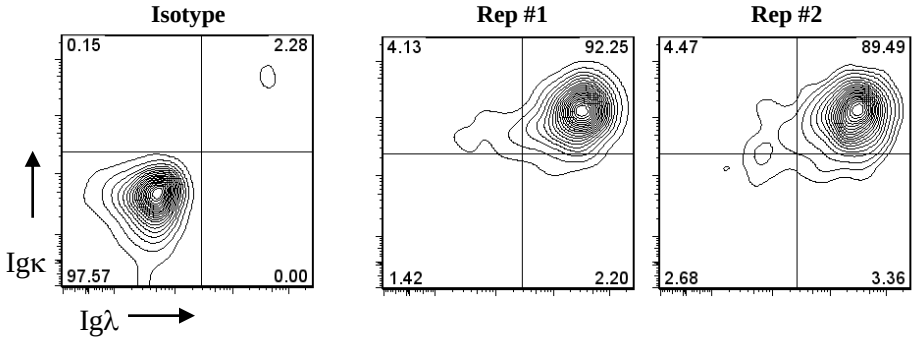


Figure 5.6 Investigating the stability of dual-light chain expression in an SLE patient. Immunoglobulin light chain expression by $CD19^+$ B cells from an SLE patient after co-culture with allogeneic $CD3^+CD4^+$ T cells and stimulation with PMA for 2 days. Cells remained positive for both Igκ and Igλ cell surface expression.

The data illustrated in this section examined the possibility that the expression of both Igκ and Igλ light chains was due to artefact generated during the immunostaining process and treatment of SLE cells prior to analysis. Treatment of SLE PMBCs with DNase and acidic wash buffers did not alter the profile of light chain expression observed. Co-culturing the cells under stimulating conditions in attempts to induce recycling of the BCR and subsequent shedding of potential immunoglobulin-containing complexes to the cell surface also did not alter the light chain expressed observed.

Previous studies have identified an unconventional population of human B cells co-expressing conventional light chains and the surrogate light chain VpreB (V-preB$^+$L$^+$). It was demonstrated that 68% of the antibodies expressed by V-preB$^+$L$^+$ cells were autoreactive; 50% were found to be genuine anti-nuclear antibodies which held additional binding complementarity to LPS, insulin and both ss- and dsDNA. In addition, these cells were found to express mRNA transcripts of RAG proteins and display skewed *IGKJ* gene usage, potentially an imprint of receptor editing (Meffre et al., 2000b; Meffre et al., 2004). V-preB$^+$L$^+$ B cell populations in peripheral blood are of very low frequency, but it was considered an important factor to eliminate in the context of the dual light chain expression profiles observed in SLE.

To investigate the possibility that expression of VpreB could contribute to possible cross-reaction of the antibody to Igλ light chains, flow cytometry analysis was performed by immunostaining the PBMCs of 6 anonymous healthy donors and 8 SLE patients for the expression of CD19, Igκ, Igλ and the surrogate light chain VpreB. No difference was observed in the cell surface expression of VpreB between healthy donors and SLE patients (Figure 5.7 A). These data however are not supported by a positive control sample, such as PBMCs isolated from fresh bone marrow, since at the time of immunostaining for VpreB such a sample was not available.

To further investigate VpreB expression using a different method, VpreB was amplified by a semi-nested PCR reaction. The primers used to amplify VpreB span an intron so therefore allowed for amplification of both the DNA amplicon and RNA transcript. If VpreB has been transcribed and expressed as protein, a smaller band of approximately 350bp would be observed since the intron will have been spliced out. The positive control used was cDNA isolated from BM cells, where detection of surrogate light chain is expected during human B cell development in the large pre-B cell. PCR amplification was performed on the cDNA of single (Igκ- *and* Igλ-expressing) FACS-sorted cells from the SLE patient

gated in Figure 5.6. VpreB transcript was only successfully amplified from the BM cDNA sample, where transcripts of this surrogate light chain would be expected (Figure 5.7 B). Bands the expected size of DNA VpreB amplicons were detected in two of the three SLE single cell cDNA samples (lanes 4 and 5) indicating the potential presence of a DNA contaminant in these samples. These data would require repeating using the correct experimental controls for the VpreB flow cytometry analysis and removal of the contaminant to be conclusive.

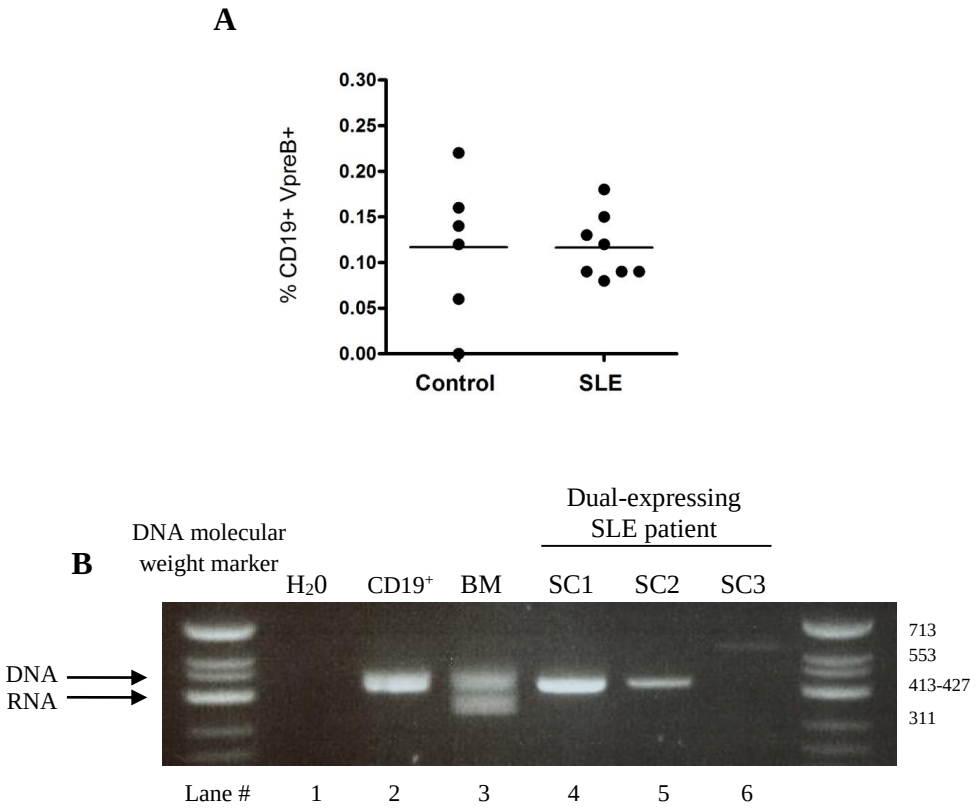


Figure 5.7 Investigating the expression of surrogate light chain VpreB by $CD19^+$ B cells in healthy controls and SLE **(A)** Flow cytometry analysis of VpreB expression by $CD19^+$ B cells. No differences were observed between healthy donors (n=6) and SLE patients (n=8). **(B)** Gel electrophoresis of PCR products after amplification of VpreB from total $CD19^+$ DNA (lane 2), cDNA isolated from BM (lane 3), and in the cDNA isolated from single dual light chain expressing $CD19^+$ cells from one SLE patient (lanes 4-6). VpreB transcript could only be detected from the BM cDNA sample.

5.2.3 cDNA analysis of single FACS-sorted SLE B cells

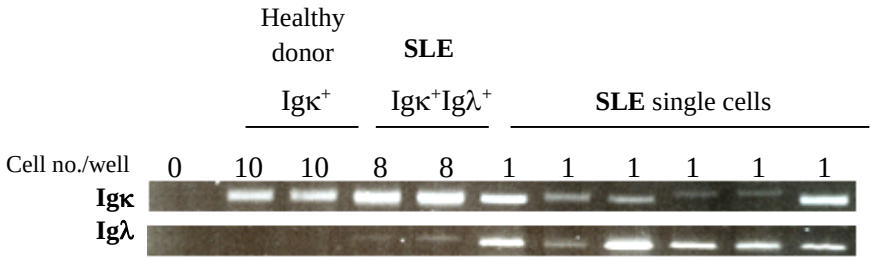
The attempts to challenge the observation that SLE B cells may frequently express both Igκ and Igλ light chains have ultimately relied on flow cytometry analysis as a readout for determination of light chain production. In order to investigate dual light chain production using a different method, dual light chain expressing $CD19^+$ B cells from a dual-expressing SLE patient (the same sample as gated in Figure 5.6) were single cell sorted into 96-well plates containing SLyRT buffer and the RNA reverse transcribed from individual cells. Single $CD19^+$ B cells were sorted into 80 wells of each plate (rows A-H, columns 1-10) where column 11 contained 0 cells per well (negative control wells) and column 12 contained 8 cells per well (positive control wells). The reverse transcribed product was then PCR amplified in two separate reactions, one containing primers to amplify VκCκ (for *IGK* transcripts) and the second to amplify VλCλ (for *IGL* transcripts). PCR primers were then used to further amplify the PCR products of round 1 in a second round.

As experimental controls, ten Igκ-expressing $CD19^+$ cells were sorted from a healthy individual into a 96-well plate containing SLyRT buffer and similarly RNA was reverse transcribed to cDNA prior to PCR amplification in two discrete reactions for the screening of *IGK* and *IGL* transcripts. The *IGL* PCR was performed on cDNA from $Ig\kappa^+$ healthy B cells to serve as the negative control alongside a PCR H_20 control. For analysis of light chain expression from the single cells, PCR for the expression of *IGK* transcripts was initially performed, and subsequent PCR amplification for *IGL* was only performed on the cDNA from wells that were positive for an *IGK* VκCκ transcript.

Igκ-expressing cells from the healthy donor expressed *IGK* but not *IGL* message. *IGK* and *IGL* transcripts of the correct band size were both detected in the wells where 8 dual-expressing SLE $CD19^+$ cells had been sorted (Figure 5.8 A).

The efficiency of *IGK* light chain transcript detection from the single cells sorted from the SLE patient was 70%, and of these, 30% were also positive for *IGL* transcript. As a result, PCR amplicons of *IGK* and *IGL* from six single dual-expressing cells were cloned and sequenced and were found to be productive and unique after analysis on the IMGT database (Figure 5.8 B).

A



B

Cell #	Productive Kappa cDNA		Productive Lambda cDNA	
1	IGKV3D-20	IGKJ4	IGLV1-40	IGLJ3
2	IGKV1-39	IGKJ1	IGLV1-40	IGLJ3
3	IGKV3-15	IGKJ4	IGLV2-14	IGLJ3
4	IGKV2-28	IGKJ2	IGLV2-11	IGLJ2
5	IGKV2-30	IGKJ2	IGLV1-44	IGLJ3
6	IGKV1-8	IGKJ2	IGLV2-8	IGLJ2

Figure 5.8 cDNA analysis of SLE dual-expressing B cells. (A) Detection of both *IGK* and *IGL* transcripts in six single FACS-sorted SLE B cells following PCR amplification (B) IMGT determination of *IGK* and *IGL* gene segment usage in each of the 6 dual light chain expressing B cells.

5.2.4 Investigating the origin and nature of dual light chain expression in SLE

Detection of both *IGK* and *IGL* rearranged transcripts within single cells of an SLE patient supported the hypothesis that the observed cell surface expression of both Igκ and Igλ light chains in some SLE patients may be a genuine feature of SLE B cells.

5.2.4.1 Studying somatic hypermutation in SLE dual-expressing B cells

Rearrangement of the KDE functionally silences unwanted rearrangements of *IGK*. This rearrangement removes *IGKC* and the intronic enhancers required for the accumulation of somatic hypermutations. Failure of KDE rearrangement could therefore result in somatic hypermutation of non-productive alleles of *IGK* that would normally have been inactivated and therefore would not undergo mutation. In order to examine the somatic mutation status of *IGK* rearrangements from dual light chain expressing SLE B cells, VκJκ rearrangements were PCR amplified from genomic DNA isolated from FACS sorted $CD19^+$ cells from two SLE patients exhibiting 87% and 58% dual light chain expression. PCR products were then cloned and sequenced, and the VκJκ rearrangements obtained were analysed using IMGT, where the productivity and mutation load could be assessed.

SLE sequences were also compared to data obtained by a previous member of the lab, Dr. Su Wen, who examined the frequency of mutations within productively and non-productively rearranged sequences from DNA of FACS sorted single lamina propria plasma cells. Studying the frequency of mutations within non-productive rearrangements reflect the cell's capacity to silence unwanted rearrangements of *IGK* by KDE rearrangement, and any mutations observed in these sequences can be assumed to be nucleotide errors introduced from the use of Taq polymerase during the PCR reaction.

Studying the mutational frequency of rearranged *IGK* light chains isolated from the gDNA of FACS-sorted LP plasma cells serves as a good positive control since these cells represent a chronically activated population. Productive VκJκ rearrangements from these cells carry a high load of mutations, whereas in the non-productive rearrangements, 67% carry no mutations (31/46) and configure with the germline Vκ and Jκ gene segments (Figure 5.9). Remaining sequences only carried few numbers of mutations, assumed to be the consequence of the errors introduced by Taq polymerase during PCR-amplification. Strikingly, non-productive VκJκ rearrangements from the pooled SLE sequences carried similar levels of mutations to their productive equivalents, where just 47% of non-productive *IGK* rearrangements carried no mutations (15/32) compared to 56% (24/43) within the productively rearranged SLE sequences.

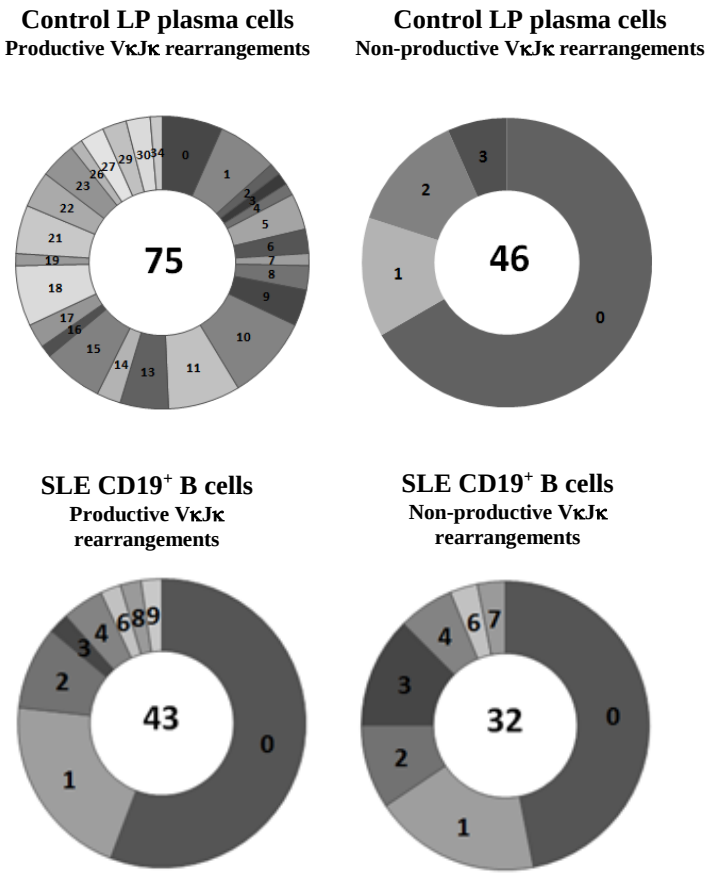


Figure 5.9 Load of somatic mutations observed within productively and non-productively rearranged VκJκ sequences isolated from the genomic DNA of single LP plasma cells from a healthy individual compared with those isolated and amplified from $CD19^+$ B cells from two SLE patients exhibiting aberrant light chain cell surface expression. The number of individual sequences analysed is in the centre of each pie chart. ***The mutational analysis from healthy intestinal LP plasma cells was performed by a former member of the Spencer Lab, Dr. Su Wen.*

5.2.4.2 Comparison of dual light chain expression by different B cell subsets in health and SLE

In order to investigate the origin and nature of dual expression seen in SLE, it was first asked whether dual expression was a feature confined to a specific B cell subset. In order to examine this, PBMCs from 10 SLE patients and 7 healthy controls were stained and

analysed by flow cytometry. The following subsets were identified; early emigrant transitional B cells ($CD19^+$ IgD^+ $CD27^-$ $CD10^{hi}$), mature but antigen-naive B cells ($CD19^+$ IgD^+ $CD27^-$ $CD10^{lo}$), memory B cells ($CD19^+$ IgD^- $CD27^+$ $CD10^-$) and activated plasmablasts, positive or negative for the production of IgA ($CD19^+$ $CD38^+$ $CD27^+$ $IgA^{hi/-}$). Cells within each B cell subset were then further divided according to their cell surface light chain expression so that cells expressing both light chains could be identified. In health, a significantly higher frequency of IgA-producing plasmablasts appeared to express both Igκ and Igλ light chains compared to non-IgA-producing plasmablasts (p=0.0379*). However, the percent frequency of dual expression across all subsets analysed is much lower than the $Ig\kappa^+$ $Ig\lambda^+$ frequencies observed in SLE (Figures 5.10 A and B). Healthy controls typically harbour approximately 2-3% $Ig\kappa^+$ $Ig\lambda^+$ $CD19^+$ B cells. In SLE however, if B cell samples exhibited dual expression of both light chains this was observed among all B cell subsets analysed and at a higher frequency than the healthy control subsets studied (Figure 5.10 B).

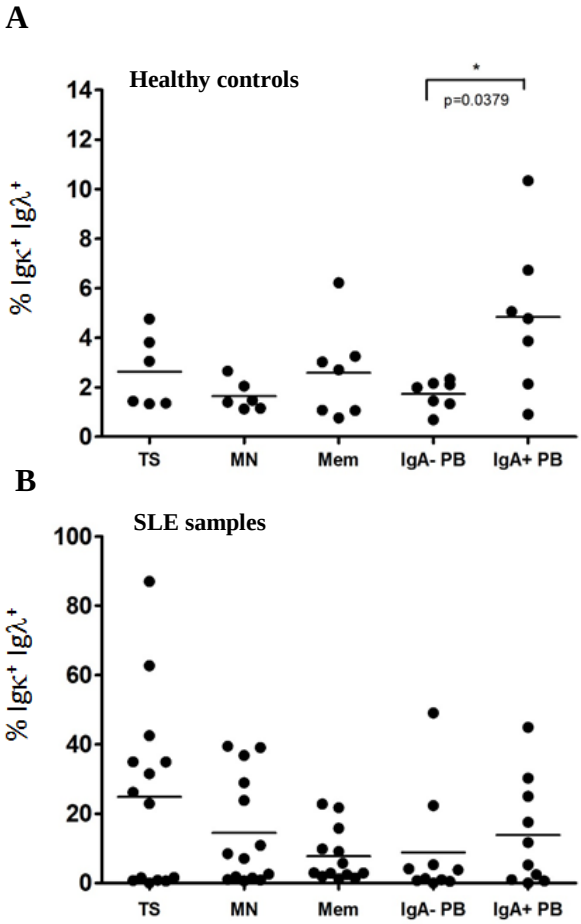


Figure 5.10 Flow cytometry analysis of the frequency of dual light chain expressing B cells in different B cell subsets; transitional (TS) B cells ($CD19^+$ IgD^+ $CD27^-$ $CD10^{hi}$), mature naive (MN) B cells ($CD19^+$ IgD^+ $CD27^-$ $CD10^{lo}$), memory (Mem) B cells ($CD19^+$ IgD^- $CD27^+$ $CD10^-$) and activated non-IgA and IgA-producing plasmablasts (PB) ($CD19^+$ $CD38^+$ $CD27^+$ $IgA^{-/hi}$) in healthy **(A)** controls and **(B)** SLE. Note the difference in scales of both A and B. IgA plasmablast B cell populations express significantly higher levels of dual Igκ and Igλ light chain surface expression than that observed in the other B cell subsets investigated, though this expression is lower than the frequencies observed in SLE. Statistical analyses were performed by Mann-Whitney tests assuming non-parametric distributions. *$p<0.05$

As a result of these data it was asked whether dual Igκ and Igλ light chain cell surface expression may be associated with a failure to rearrange the KDE, since Chapter 4 demonstrated reduced KDE rearrangement status in Igλ-expressing SLE B cells compared

to equivalent Igλ-expressing B cell from healthy individuals. It was observed that a statistically significant negative correlation (p=0.0072) existed between the proportion of $CD19^+$ $Ig\kappa^+Ig\lambda^+$ B cells and KDE rearrangement activity in SLE (Figure 5.11). This was not apparent following similar analysis of healthy subjects.

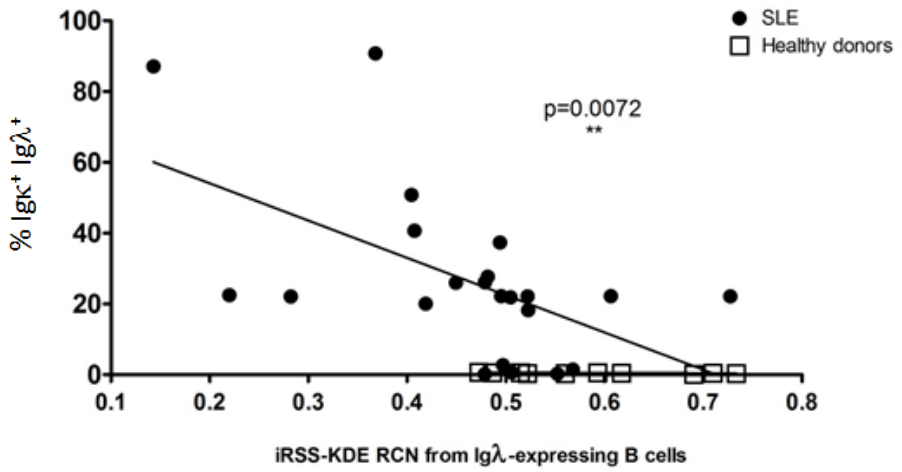


Figure 5.11 Negative correlation between rearrangement status of the KDE and cell surface light chain expression of both Igκ and Igλ light chains by $CD19^+$ B cells in SLE. (Data is non-parametrically distributed.)

5.2.6 Free light chains in SLE patient serum

Quantification of free light chains in human serum is a frequently used immunoassay where increased detection of free light chains have been reported to be a useful biomarker in several autoimmune diseases. In order to ask whether dual light chain expression by SLE B cells may be associated with this phenomenon, serum samples from several SLE patients already analysed by flow cytometry were examined for the levels of free Igκ and Igλ light chains. This was a collaboration with The Binding Site Group Ltd, Birmingham. Interestingly, there was a significant correlation between the cell surface expression of both

Igκ and Igλ light chains and the levels of free Igκ or Igλ light chains in the serum of the same patients (Figure 5.12 A and B).

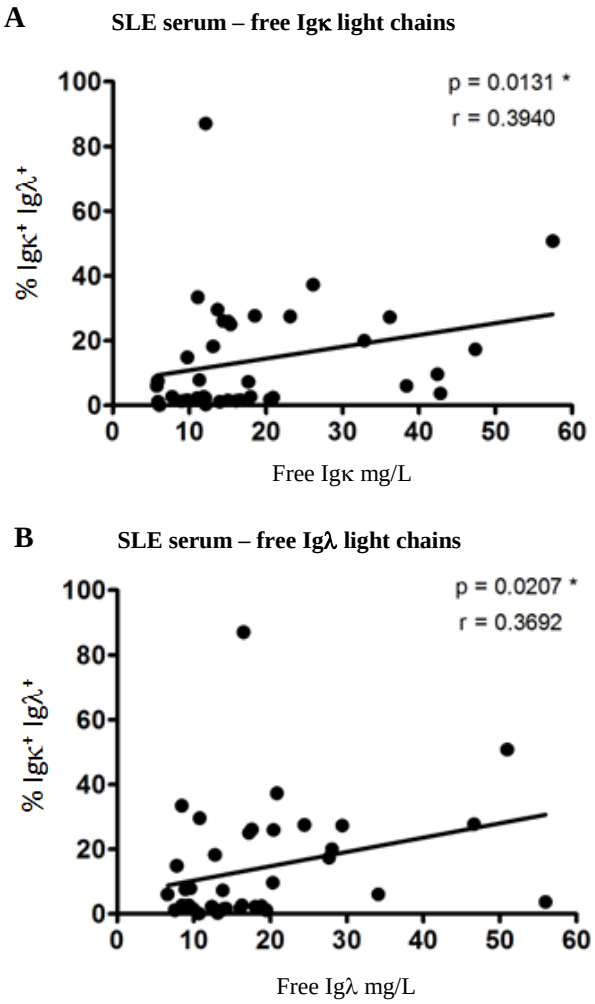


Figure 5.12 Correlation between the frequency of dual light chain expression by $CD19^+$ SLE B cells and the concentration of **(A)** free Igκ and **(B)** free Igλ light chains in the serum of the same patients. SLE patients with higher quantities of both serum free Igκ and Igλ light chains tended to harbour larger populations of $CD19^+$ peripheral B cells expressing both Igκ and Igλ light chains on the cell surface. Statistical analysis performed by Mann-Whitney test assuming non-parametric distributions. *$p<0.05$

5.2.7 Summary of experimental findings

- **Is a tendency to express both Igκ and Igλ light chains by some B cells a rare or frequent observation?**

  Outcome: Flow cytometry analysis revealed that approximately half of all SLE patients examined in this study carried $CD19^+$ B cells with clear dual-light chain expression populations.

- **Is apparent dual Igκ and Igλ light chain expression a genuine feature of SLE $CD19^+$ B cells, or is cytophilic binding of Igs, or Ig-ICs bound to the surface of SLE B cells responsible for interfering with immunostaining prior to flow cytometry analysis?**

  Outcome: The proportion of SLE $CD19^+$ B cells expressing both Igκ and Igλ light chains does not differ following DNase treatment or acidic washing prior to immunostaining.

- **Is dual light chain cell surface expression stable?**

  Outcome: Dual-light chain expressing $CD19^+$ B cells FACS sorted from an SLE patient maintained expression of both light chains following a 2 day co-culture with allogeneic $CD3^+CD4^+$ T cells under stimulatory conditions.

- **Could VpreB surrogate light chain expression account for the apparent dual expression of light chains?**

  Outcome: VpreB transcript could not be detected by flow cytometry analysis of SLE PBMCs, or by PCR of cDNA isolated from single SLE dual-expressing B cells.

- **Can rearranged transcripts of *IGK* and *IGL* be detected in the cDNA isolated from single cells of a dual Igκ$^+$Igλ$^+$ SLE patient?**

  Outcome: Unique and productively rearranged *IGK* and *IGL* sequences (Vκ-Cκ) were obtained following the cloning of PCR products amplified from the cDNA of single FACS-sorted CD19$^+$Igκ$^+$ Igλ$^+$ B cells from an SLE patient.

- **Is dual light chain expression a feature of one specific B cell subset?**

  Outcome: Flow cytometry analysis of different B cell subsets revealed that when dual Igκ and Igλ expression was a feature of an individual SLE patient, it was not confined to a specific B cell subset.

- **Is there a correlation between dual light chain expression and KDE rearrangement activity in SLE?**

  Outcome: A significant negative correlation was observed between the frequency of B cells expressing Igκ and Igλ light chains, and the frequency of KDE rearrangements in the gDNA of Igλ–expressing SLE B cells.

- **Are serum free light chains associated with dual light chain expression by SLE B cells?**

  Outcome: There was a positive correlation between the levels of free light chains in the serum of SLE patients and the proportion of B cells expressing both Igκ and Igλ light chains.

5.3 Discussion

The observation that KDE rearrangements are inefficient in SLE compared to health in the previous chapter could suggest that a failure to silence rearranged but unwanted alleles of *IGK* may allow for their cell surface expression by some B cells. The global consequence of a failure to rearrange the KDE in populations of B cells could therefore result in a failure to obey isotypic exclusion, and could be accompanied by the observation of dual light chain expression by single B cells. Failure of light chain isotypic exclusion in the context of a failure to rearrange the KDE could include the expression of two different Igκ light chains simultaneously by an Igκ-producing B cell, or the expression of one or more Igκ light chains and an Igλ light chain by Igλ-producing B cells.

For the expression of more than one light chain to be biologically plausible, the rearranged alleles of *IGK* unsilenced by KDE rearrangement would have to be in the correct genetic reading frame in order to be transcribed and expressed on the cell surface. As described in Chapter 3, analysis of the productive forbidden rearranged *IGK* light chain sequences obtained from the high-throughput sequencing analysis from the genomic DNA of Igλ-expressing B cells revealed that over 20% of these *IGK* rearrangements in both healthy individuals and SLE patients were productively rearranged (Figure 3.3 B). If these forbidden productive rearrangements of *IGK* do not carry rearrangements of the KDE, the expression of these Igκ light chains that were previously selected against could be conceivable. These data suggest that failure to KDE-inactivate rearrangements of *IGK* light chains that are unwanted may subsequently allow for their transcription, translation and subsequent cell surface expression by SLE B cells.

Several methods were used to challenge the dual expression and apparent allelic inclusion initially observed during flow cytometry analysis of SLE PBMCs. Treatment of the pre-

immunostained PBMCs with DNase I, washing the pre-immunostained PBMCs with acidic wash buffers, or co-culturing sorted Igκ$^+$ Igλ$^+$ cells short-term with PMA did not alter the profile of light chain expression exhibited by specific SLE samples after immunostaining. SLE PBMCs were also shown to express similar levels of cell surface VpreB expression as healthy controls, however a caveat of these data is that no positive control was available for use at the time of the immunostaining this surrogate light chain. A sample known to express cell surface VpreB, such as PBMCs isolated from fresh bone marrow, would be required if this experiment were to be repeated.

Unique *IGK* and *IGL* (V_L-C_L) rearranged sequences were obtained from the cDNA of single dual-expressing B cells from an SLE patient, suggesting genuine expression and production of two light chains is a feature of this SLE CD19$^+$ B cells. A caveat of these data lies in the lack of a control for testing the *IGL* primers; the primers used for amplification of *IGK* and *IGL* transcripts were tested using cDNA isolated from 10 Igκ- producing B cells from a healthy donor (the *IGL* primers were used on this cDNA to test for specificity to *IGK*). If this experiment were to be repeated, 10 Igλ$^+$ cells should be sorted from the same healthy individual to demonstrate that the primers detecting *IGL* transcripts are also just specific for *IGL* amplification and not *IGK*. Nevertheless, detection of *IGL* transcripts was achieved from the SLE single sorted B cells following sequence analysis using IMGT.

The hypothesis that allelic exclusion is not intact in a large proportion of the SLE patients used in this study is supported with reference to the quantification of KDE rearrangement frequencies in SLE compared to health described in Chapter 4. Populations of Igλ-expressing B cells in SLE exhibited significantly lower frequencies of KDE rearrangements compared to controls. As shown in Figure 5.11 of this chapter, there is significant correlation between SLE CD19$^+$ B cells with the lowest levels of KDE rearrangement activity to appear to express both Igκ and Igλ light chains on the cell surface. Dual

expression of both light chains by SLE B cells appears to not be confined to one specific B cell subset, which could suggest that failure to allelically exclude light chains is a phenotype that is not lost during B cell maturation since it is genetically determined during early B cell development.

In further agreement within a defective KDE rearrangement status in SLE, somatic mutations were observed with the non-productive B cell repertoire of two SLE patients at equivalent frequencies to those observed in the productive VκJκ sequences from the same patients; alleles which should have lost their intronic enhancers following KDE rearrangement. A caveat of this sequence analysis is that a more relevant controls could have been used. VκJκ rearrangements could have been amplified from gDNA isolated from $CD19^+$ healthy control B cells and also from the gDNA isolated from $CD19^+$ SLE B cells from a patient that did not exhibit dual expression of both light chains. However, the equivalent frequencies of mutations observed in the productive and non-productive *IGK* rearranged sequences from $CD19^+$ populations of B cells from two SLE patients support the hypothesis that failure to rearrange the KDE is a feature of SLE.

Failure to isotypically exclude light chains has been reported in healthy individuals and mice. An early study described the generation of a human pre-B cell line by EBV transformation, named 30.30, which co-expressed both Igκ and Igλ light chains on 49% of all cells during flow cytometry analysis (Pauza et al., 1993). These cells carried two rearrangements of *IGK*, unsilenced by the KDE, and single rearrangements of *IGL* and *IGH*. The phenotype of this dual light chain expressing population after FACS sorting was not stable, as after several weeks in culture cells became a heterogeneous pool of both $Ig\kappa^+$ $Ig\lambda^+$ and single $Ig\kappa^+$ cells (Pauza et al., 1993).

B cell Igκ light chain allelic inclusion has previously been observed, and experiments involving the introduction of a human *IGKC* region gene to be expressed alongside mouse *IGKC* genes have demonstrated that B cells harbouring both Igκ light chains were found to be responsive to BCR stimulation with anti-IgM in vitro and BCR aggregation in vivo (Velez et al., 2007). Several groups have observed a genuine population of B cells expressing two different light chains and a tendency to exhibit autoreactivity in both mice and humans (Fournier et al., 2012; Giachino et al.,1995; Pauza et al., 1993). In addition, a small population of Igκ- and Igλ- producing $CD19^{+}$ B cells following FACS and EBV-immortalisation were characterised from five dual-light chain expressing B cell clones. It was revealed that these cells expressed both Igκ and Igλ chains on the same Ig molecule, expressed IgM and three of the five clones carried modest numbers of somatic mutations in Vκ-Jκ, suggesting that they were isolated from an antigen-selected B cell population (Giachino et al., 1995).

A recent study crossed IGK^{h} mice that carry gene targeted human *IGKC* allele within a wild-type Ig repertoire with lupus-prone MRL-*fas*$^{lpr/lpr}$ mice to allow for study of in the context of a self-intolerant autoimmune disease state (Fournier et al., 2012). These mice exhibit defective receptor editing and reduced tolerance induction (Li et al., 2002a; Panigrahi et al., 2008). The dual Igκ-producing B cells from these mice exhibited autoreactivity and their frequency increased with age and disease severity. They were also found to be contained within the plasmablast memory B cell pool and were capable of secreting autoantibodies (Fournier et al., 2012). More recent work has demonstrated that usage of both *IGK* and *IGL* light chain gene segments by B cells has a profound effect on the maintenance and breakage of B cell tolerance in animal models of autoimmunity (Andrews et al., 2013).

Chapter 6
Overview and potential future directions

The high-throughput sequencing analysis of Chapter 3 studied the profile of *IGK* rearrangements in the DNA and cDNA of mature naïve B cells isolated from the peripheral blood of three healthy donors and three SLE patients. The profile of *IGK* rearrangements in the genomic DNA of mature naïve B cells bears the hallmarks of *IGKV* segment selection during B cell development and this is equivalent in health and SLE. Selection both for and against individual segments was apparent and consistent with an earlier analysis of genomic DNA from single IgM-expressing B cells. It is unclear what aspect of functionality is selected for, but the equivalence of this process in health and SLE however demonstrates that the mechanisms that identify these segments have occurred at the pre-B stage and the subsequent cell-fate decision to either continue to mature or move to rearrange at the *IGL* locus is the same in SLE and health with respect to segment selection.

The biases in the *IGK* repertoire in SLE were apparent by analysis of *IGK* gene expression at the cDNA level. This suggests that the mechanisms regulating the expression of *IGK* rearrangements are responsible for the biased light chain repertoire in SLE and not the segment selection process *per se*. This study identified that *IGKV5-2* is selected against during development at the point where cells make the fate decision to express *IGK* or to move on and rearrange *IGL*. In contrast, *IGKV4-1* appears not to be selected against at this point, though both segments have decreased representation in the expressed repertoire in health. This suggests that multiple mechanisms are involved in identifying and regulating the *IGK* gene expression, presumably including those that initiate and regulate KDE activity. These may act at different times during B cell development. Their activity is essential because most B cells in the blood contain multiple light chain gene rearrangements. Failure to effectively regulate can lead to aberrant expression, potentially also allelic inclusion and a less stringently regulated light chain repertoire. The caveat of these data described in Chapter 3 lies in the unequal numbers of sequences obtained from each B cell subset and across all individual samples, though time and funding did not permit

the repeat of such experiments, which could have ensured equal numbers of cells sorted from each individual. Nevertheless, these data have contributed to the identification of a potential origin of the light chain biases observed in SLE.

Chapter 4 explored the rearrangement activity of the kappa deleting element, which in part plays a role in regulating *IGK* expression through the silencing of *IGK* rearranged but unwanted alleles. The data in this chapter demonstrated that populations of SLE $CD19^+$ B cells carried reduced rearrangement frequencies of the KDE compared to healthy controls, which could account for the emergence of *IGK* rearrangements involving *IGKV4-1* and *IGKV5-2* in the expressed repertoire in SLE where they are rarely observed in the expressed repertoire in health, as observed in Chapter 3. Support for the observation of reduced KDE rearrangement activity by SLE $CD19^+$ B cells was demonstrated by sequence analysis of somatic mutations within non-productively rearranged *IGK* alleles of $CD19^+$ B cells from 2 SLE patients. The equivalent frequencies of mutations observed in both the non-productive and productive groups of sequences indicates that the *IGKC* gene segment and enhancers flanking it may not have been removed in some instances. This would permit the acquisition of mutations by some *IGK* rearrangements that theoretically should have been destined for KDE inactivation.

Chapter 5 investigated the potential for populations of B cells in some SLE patients to exhibit altered light chain cell surface expression profiles, as observed in previous chapters. Approximately half of all SLE patients studied in this investigation appeared to harbour populations of $CD19^+$ B cells expressing both Igκ and Igλ light chains on their cell surface following flow cytometry analysis. Experiments to challenge the validity of these observations did not disprove the flow cytometry data. Productive transcripts of both *IGK* and *IGL* were detected in an individual SLE patient that appeared to exhibit dual Igκ and Igλ cell surface expression on approximately 90% of their peripheral B cells. Perhaps most interestingly, the frequency of dual light chain-expressing B cell populations in SLE

negatively correlated with the frequency of KDE rearrangements quantified in Igλ-expressing B cell gDNA. Dual expression of both Igκ and Igλ light chains also appeared to positively correlate with levels of serum free Igκ and Igλ light chains, a marker typically associated with autoimmune disease states in humans. These findings may suggest that KDE rearrangement activity alongside analysis of cell surface light chain expression would be a useful tool in predicting SLE disease. A key point to note is that a history of nephritis, both active and non-active at the time the SLE blood was taken is significantly associated with the expression of both Igκ and Igλ light chains, the latter of which negatively correlates with KDE rearrangement frequency in population of SLE B cells. What would be fundamental to next address would be identification of the signalling pathway or pathways that initiate rearrangement of the KDE, and potential proteins that may interact and bind the KDE. These questions could be investigated using a ChIP approach. If this could be achieved any differences in the protein machinery governing KDE rearrangement could be investigated.

The work presented in this thesis proposes that defective regulation of light chain gene expression in SLE, which in part may be mediated by a failure to rearrange the KDE, is relevant to SLE pathogenesis by aberrantly permitting the 'breakthrough' expression of segments that have been counter-selected. Individual B cells that allelically include Ig light chains may be capable of recognising and responding to more than one foreign antigen via their BCRs. The presence of these cells in the peripheral setting of an SLE autoimmune T cell repertoire could serve to propagate and amplify autoimmunity. These findings may also contribute to understanding the origin of the self-reactive and hyper-responsive nature of SLE B cells and the breakdown in immune tolerance associated with SLE pathogenesis.

Appendices

Appendix 1 *Mature naive B cell population frequencies following sorting*

Sample status	**Sample ID**	**CD19⁺IgD⁺CD27⁻**	
		Igκ+	Igλ+
Healthy donors	**H01**	54,000	54,000
	H02	100,000	100,000
	H03	115,000	115,000
SLE patients	**SLE 1**	42,750	42,750
	SLE 2	15,000	15,000
	SLE 3	180,000	180,000

Igκ- and Igλ-expressing sorted B cell populations were divided into two so approximately half of each total population was used for DNA isolation and RNA isolation.

Sequence frequencies following high-throughput sequencing analysis

Healthy controls (n=3)

CD19⁺IgD⁺ CD27⁻	**DNA / cDNA**	**PCR**	**Experimental name**	**Productive**				**Non-Productive**			
				HC 1	**HC 2**	**HC 3**	**TOTAL**	**HC 1**	**HC 2**	**HC 3**	**TOTAL**
Igκ+	cDNA	Vκ-Cκ	Used	144	295	707	**1146**	NA	NA	NA	0
Igκ+	gDNA	Vκ-Jκ	All rearr.	173	542	451	**1166**	103	460	305	**868**
Igλ+	gDNA	Vκ-Jκ	Not used	295	308	10	**613**	955	961	27	**1943**
										Total	**5736**

SLE patients (n=3)

CD19⁺IgD⁺ CD27⁻	**DNA / cDNA**	**PCR**	**Experimental name**	**Productive**				**Non-Productive**			
				SLE 1	**SLE 2**	**SLE 3**	**TOTAL**	**SLE 1**	**SLE 2**	**SLE 3**	**TOTAL**
Igκ+	cDNA	Vκ-Cκ	Used	175	1219	499	**1893**	NA	NA	NA	0
Igκ+	gDNA	Vκ-Jκ	All rearr.	113	296	340	**749**	45	262	287	**594**
Igλ+	gDNA	Vκ-Jκ	Not used	234	70	6	**310**	878	254	20	**1152**
										Total	**4698**

Appendix 2 *SLE patient clinical information*

Sample ID	Sex	Age	Ethn	Disease duration	% CD19+ Igκ+ Igλ+	Treatment
Lu-009	F	48	AfC	7	0.26	MTX, HCQ, Pred
Lu-011	F	42	Cauc	22	0.13	MTX, HCQ, Pred
Lu-014	F	57	Cauc	10	0.64	MMF, Pred
Lu-015	F	23	Cauc	8	20.02	HCQ
Lu-016	F	49	AfC	17	18.23	HCQ
Lu-017	F	45	Cauc	16	26.06	MMF, HCQ, Mepacrine
Lu-018	F	65	Cauc	8	1.45	none
Lu-022	F	73	Cauc	25	27.67	HCQ
Lu-023	F	45	Cauc	15	2.71	HCQ
Lu-024	F	45	Af	6	37.32	HCQ, Pred
Lu-025	F	40	Cauc	14	87.06	HCQ, Pred
Lu-026	F	53	Cauc	18	1.18	none
Lu-031	F	21	AfC	3	58.15	AZA, HCQ
Lu-039	F	40	Asian	16	50.8	AZA, HCQ, Pred
Lu-040	F	35	Orien	13	25.97	AZA, Pred
Lu-041	F	51	Cauc	9	2.31	MMF, HCQ, Pred
Lu-042	F	63	Cauc	20	0.82	AZA, Pred
Lu-043	F	46	Af	5	90.73	MMF, HCQ, Pred
Lu-044	F	54	Af	4	4.53	HCQ, Pred
Lu-045	F	40	Af	12	14.85	MMF, HCQ, Pred
Lu-046	F	23	Cauc	2	1.18	MMF, HCQ, Pred
Lu-047	F	59	Cauc	4	1.48	Leflunomide, HCQ
Lu-048	F	32	Cauc	3	2.68	HCQ, Pred
Lu-049	F	51	Cauc	4	0.31	none
Lu-056	F	58	Cauc	39	29.6	none
Lu-061	F	31	Af	5	2.7	RTX 2008 2009 2011 2012 HCQ, Thalidomide, Pred
Lu-066	F	44	Cauc	18	3.71	MMF, HCQ, Pred
Lu-067	F	52	AfC	16	27.3	RTX 2012 HCQ, Pred
Lu-069	F	47	Af	20	27.5	CYC IV HCQ Pred
Lu-070	M	62	Cauc	20	25	HCQ, Pred
Lu-071	F	67	AfC	1	1.8	RTX 2012 HCQ, Pred
Lu-073	F	33	Af	9	7.84	HCQ, Pred
Lu-075	F	56	Cauc	12	7.6	RTX 2009 2012 AZA HCQ Pred
Lu-076	F	34	Af	5	6.8	RTX 2010 2012
Lu-077	F	51	Cauc	21	26	HCQ, Pred
Lu-078	F	23	Oriental	1	2.2	HCQ, Pred
Lu-079	F	54	Asian	40	17.3	HCQ, Pred
Lu-080	F	57	Cauc	14	7.3	HCQ
Lu-081	F	36	Cauc	14	2	RTX 2012 MMF HCQ Pred
Lu-086	F	60	Cauc	12	1.5	RTX 2012 MTX HCQ Pred
Lu-092	F	17	Af	10	6	RTX 2010 2012 Pred
Lu-095	F	32	Af	5	2.4	RTX 2012 MMF HCQ Pred
Lu-104	F	40	Cauc	20	9.62	RTX 2009 2012 AZA HCQ Pred
Lu-108	F	26	Af	7	26.3	RTX 2012 MMF HCQ Pred
LCLu-008	F	55	Cauc	Unknown	0.26	Statins
LCLu-118	F	62	Oriental	13	22.15	MMF, lisinopril, aspirin, SIM, calcichew
LCLu-119	F	47	AfC	13	22.07	Pred, AZA, LOS, RAN, Plaqenil
LCLu-120	F	38	Cauc	16	22.17	AZA, Pred, warfarin, aspirin, adalat, HCQ, rabeprazole, co-codamol
LCLu-121	F	41	Cauc	7	10.68	CAN, HCQ, MTX, aspirin, NTP, TRA, OMP, calcichew, alendronate acid, AML, co-dydrmol
LCLu-122	F	37	Asian	12	21.84	not listed
LCLu-124	F	25	Asian	13	40.7	Pred, myfortic,warfarin, VAL, ranitidine, HCQ, vitamin d, calcium
LCLu-125	F	24	Oriental	5	22.49	Pred, AZA, HCQ, famotidine, SIM, enalapril
LCLu-126	F	46	AfC	21	0.78	not listed
LCLu-127	F	41	AfC	18	1.22	not listed
LCLu-128	F	56	AfC	3	1.71	Pred, HCQ, adcal-3-calcium, ramipril, MMF, alendronic acid
LCLu-129	F	39	Asian	30	3.12	Pred, AZA, HCQ

Sample ID	dsDNA	ANA	ENA	Ro	La	Sm	RNP	C1QAb	aPL	C3	C4	Malar rash	Photo-sens	SCLE	Discoid	Arthritis	Mouth ulcers	Lupus Nephritis	CNS	CVS/RS	Haem
Lu-009	Y	Y	Y	N	N	Y	Y	N	N	0.71	0.13	N	Y	Y	N	Y	Y	LN III+V	N	N	Y
Lu-011	N	Y	Y	Y	Y	N	N	N	Y	1.41	0.23	Y	Y	Y	N	Y	Y	N	N	Y	N
Lu-014	Y	Y	Y	Y	N	N	N	N	N	1.35	0.16	N	Y	Y	N	Y	Y	LN IV+V	N	N	Y
Lu-015	Y	Y	Y	Y	N	N	Y	N	Y	1.05	0.24	Y	Y	Y	Y	Y	Y	N	N	N	Y
Lu-016	N	Y	Y	Y	N	N	Y	N	N	1.04	0.22	N	Y	Y	N	Y	Y	N	N	N	N
Lu-017	N	Y	Y	Y	N	Y	N	N	Y	0.81	0.18	N	Y	Y	N	Y	Y	LN V	N	N	Y
Lu-018	N	Y	Y	Y	N	Y	Y	N	N	0.97	0.19	Y	Y	N	N	Y	Y	N	N	N	Y
Lu-022	Y	Y	Y	Y	N	N	N	N	N	1.09	0.15	Y	Y	N	N	Y	Y	N	N	N	Y
Lu-023	Y	Y	Y	Y	Y	N	N	N	Y	0.93	0.15	Y	Y	N	N	Y	Y	N	N	N	Y
Lu-024	Y	Y	N	Y	Y	N	N	N	N	0.77	0.12	N	N	N	N	Y	Y	LN III+V	N	N	Y
Lu-025	Y	Y	Y	Y	N	N	N	N	Y	1.22	0.14	Y	Y	N	N	N	Y	N	N	N	Y
Lu-026	Y	Y	Y	Y	Y	N	N	N	N	1.36	0.14	Y	Y	N	N	N	Y	N	N	N	Y
Lu-031	Y	Y	Y	Y	N	N	N	N	N	0.75	0.08	Y	Y	Y	N	Y	Y	LN IV+V	N	N	Y
Lu-039	Y	Y	N	N	N	N	N	N	N	0.91	0.06	N	N	N	N	Y	N	LN V	N	Y	Y
Lu-040	N	Y	N	Y	Y	N	N	N	N	0.67	0.1	Y	Y	N	N	N	N	LN IV	N	N	N
Lu-041	Y	Y	N	Y	N	Y	Y	N	N	0.65	0.06	Y	N	N	N	N	N	N	Y	N	N
Lu-042	N	Y	N	N	N	N	N	N	N	1.49	0.35	Y	Y	N	N	Y	N	N	N	N	N
Lu-043	Y	Y	N	Y	N	Y	Y	N	N	1.65	0.19	N	N	N	N	Y	N	LN III	N	N	N
Lu-044	N	Y	N	N	N	N	Y	N	N			N	N	N	N	Y	N	N	N	Y	Y
Lu-045	N	Y	N	N	N	N	N	N	Y	1.08	0.45	N	N	N	N	N	N	LN II+IV	N	N	N
Lu-046	Y	Y	N	Y	Y	N	N	N	N	1.02	0.21	N	N	N	N	N	N	LN III	N	N	N
Lu-047	N	Y	N	N	N	N	N	N	N	1.49	0.3	N	N	N	N	Y	Y	N	N	N	N
Lu-048	N	Y	N	N	N	N	N	N	N	0.86	0.16	N	N	N	N	Y	N	N	N	N	Y
Lu-049	N	Y	N	Y	N	N	N	N	N	0.94	0.22	N	Y	Y	N	N	Y	N	N	N	N
Lu-056	N	Y	N	N	N	N	N	N	Y	1.05	0.2	N	Y	N	N	Y	Y	N	Y	N	N
Lu-061	Y	Y	N	N	N	N	N	N	N	0.88	0.14	N	Y	N	Y	Y	Y	N	N	N	N
Lu-066	Y	Y	Y	N	N	N	Y	N	N	0.76	0.14	Y	Y	Y	N	Y	N	LN II	Y	Y	Y
Lu-067	Y	Y	Y	N	N	N	N	Y	N	1.22	0.32	N	N	N	N	Y	Y	LN III+V	N	N	Y
Lu-069	Y	Y	Y	Y	N	Y	Y	Y	N	0.72	0.07	N	N	N	Y	Y	Y	LN IV+V	N	N	Y
Lu-070	N	Y	Y	Y	Y	N	N	N	N	1.66	0.29	N	N	N	N	Y	Y	N	N	Y	N

Sample ID	dsDNA	ANA	ENA	Ro	La	Sm	RNP	C1QAb	aPL	C3	C4	Malar rash	Photo-sens	SCLE	Discoid	Arthritis	Mouth ulcers	Lupus Nephritis	CNS	CVS/RS	Haem
Lu-071	N	Y	Y	N	N	Y	Y	N	N	1.08	0.36	N	Y	N	N	Y	Y	N	N	Y	Y
Lu-073	N	Y	Y	Y	N	Y	N	N	N	1.5	0.45	N	N	N	N	Y	Y	N	N	N	Y
Lu-075	N	N	N	N	N	N	N	N	N	1.38	0.23	N	Y	N	N	Y	Y	N	Y	Y	N
Lu-076	Y	Y	Y	N	N	Y	Y	Y	N	0.55	0.06	N	Y	N	Y	Y	Y	N	Y	N	Y
Lu-077	Y	Y	N	N	N	N	N	N	Y	0.58	0.13	Y	Y	Y	N	Y	N	N	N	N	Y
Lu-078	Y	Y	Y	Y	N	N	N	N	N	1.09	0.2	Y	Y	N	N	Y	Y	N	N	N	Y
Lu-079	N	Y	N	N	N	N	N	N	Y	1.23	0.24	N	N	N	N	N	N	LN II	N	N	N
Lu-080	N	Y	N	N	N	N	N	N	Y	1.37	0.23	Y	Y	N	N	Y	Y	LN V	N	N	Y
Lu-081	N	Y	Y	Y	N	Y	Y	N	N	1.14	0.16	Y	Y	N	N	Y	Y	N	Y	Y	Y
Lu-086	N	Y	Y	Y	Y	N	N	N	N	1.33	0.24	Y	Y	N	N	Y	Y	N	N	N	Y
Lu-092	Y	Y	Y	N	N	Y	Y	N	N	0.48	0.05	Y	Y	N	N	Y	Y	LN III+V	N	N	Y
Lu-095	Y	Y	Y	Y	N	Y	Y	N	Y	1.01	0.44	Y	Y	N	N	Y	Y	LN IV+V	N	N	N
Lu-104	Y	Y	N	N	N	N	N	N	N	0.91	0.16	Y	Y	N	N	Y	Y	LN V	N	Y	N
Lu-108	Y	Y	Y	N	N	N	N	N	N	1.11	0.44	N	N	N	N	Y	Y	LN III+V	N	N	Y
LCLu-008	N	N		N	N	N	N						N			Y	Y	N	N	N	
LCLu-118	N	Y		Y	N	N	N						N			Y	Y	Y	N	RAY	
LCLu-119	Y	Y		N	N	Y	Y						Y			Y	Y	Y	Y	RAY	
LCLu-120	N				N								N			Y	N	Y	Y	RAY	
LCLu-121	Y	Y		Y	N	Y	Y						Y			N	Y	Y	N	N	
LCLu-122	N	Y		Y	N	N	N						N			Y	Y	Y	Y	N	
LCLu-124	N			N	N	N	N						N			Y	Y	Y	N	RAY	
LCLu-125	Y	Y		Y	N	Y	Y						Y			Y	Y	Y	Y	RAY	
LCLu-126	N	Y		N	N	N	Y						Y			Y	Y	Y	Y	RAY	
LCLu-127	Y	Y		N	Y	N	N						N			Y	Y	N	N	RAY	
LCLu-128	Y			Y	N	Y	Y						Y			N	N	Y	Y	RAY	
LCLu-129	Y	Y		N	N	N	N						N			N	N	Y	Y	N	

References

Abelson, A.K., A.M. Delgado-Vega, S.V. Kozyrev, E. Sanchez, R. Velazquez-Cruz, N. Eriksson, J. Wojcik, M.V. Linga Reddy, G. Lima, S. D'Alfonso, S. Migliaresi, V. Baca, L. Orozco, T. Witte, N. Ortego-Centeno, H. Abderrahim, B.A. Pons-Estel, C. Gutierrez, A. Suarez, M.F. Gonzalez-Escribano, J. Martin, and M.E. Alarcon-Riquelme. 2009. STAT4 associates with systemic lupus erythematosus through two independent effects that correlate with gene expression and act additively with IRF5 to increase risk. *Annals of the rheumatic diseases* 68:1746-1753.

Abulayha, A.M., S.A. Tabal, E.I. Shawesh, M.A. Elbasir, A.S. Elbanani, Y.M. Lamami, and A. Bredan. 2010. Depletion of peripheral blood B cells with Rituximab and phenotype characterization of the recovering population in a patient with follicular lymphoma. *Leukemia research* 34:307-311.

Aggarwal, R., W. Sequeira, R. Kokebie, R.A. Mikolaitis, L. Fogg, A. Finnegan, A. Plaas, J.A. Block, and M. Jolly. 2011. Serum free light chains as biomarkers for systemic lupus erythematosus disease activity. *Arthritis care & research* 63:891-898.

Aguilera, I., J. Melero, A. Nunez-Roldan, and B. Sanchez. 2001. Molecular structure of eight human autoreactive monoclonal antibodies. *Immunology* 102:273-280.

Akhavan, P.S., J. Su, W. Lou, D.D. Gladman, M.B. Urowitz, and P.R. Fortin. 2013. The Early Protective Effect of Hydroxychloroquine on the Risk of Cumulative Damage in Patients with Systemic Lupus Erythematosus. *The Journal of rheumatology* 40:831-841.

Alba, P., L. Bento, M.J. Cuadrado, Y. Karim, M.F. Tungekar, I. Abbs, M.A. Khamashta, D. D'Cruz, and G.R. Hughes. 2003. Anti-dsDNA, anti-Sm antibodies, and the lupus anticoagulant: significant factors associated with lupus nephritis. *Annals of the rheumatic diseases* 62:556-560.

Allman, D., R.C. Lindsley, W. DeMuth, K. Rudd, S.A. Shinton, and R.R. Hardy. 2001. Resolution of three nonproliferative immature splenic B cell subsets reveals multiple selection points during peripheral B cell maturation. *J Immunol* 167:6834-6840.

Alt, F., N. Rosenberg, S. Lewis, E. Thomas, and D. Baltimore. 1981. Organization and reorganization of immunoglobulin genes in A-MULV-transformed cells: rearrangement of heavy but not light chain genes. *Cell* 27:381-390.

Alt, F.W., V. Enea, A.L. Bothwell, and D. Baltimore. 1980. Activity of multiple light chain genes in murine myeloma cells producing a single, functional light chain. *Cell* 21:1-12.

Alt, F.W., G.D. Yancopoulos, T.K. Blackwell, C. Wood, E. Thomas, M. Boss, R. Coffman, N. Rosenberg, S. Tonegawa, and D. Baltimore. 1984. Ordered rearrangement of immunoglobulin heavy chain variable region segments. *The EMBO journal* 3:1209-1219.

Alunno, A., E. Bartoloni, O. Bistoni, G. Nocentini, S. Ronchetti, S. Caterbi, V. Valentini, C. Riccardi, and R. Gerli. 2012. Balance between regulatory T and Th17 cells in systemic lupus erythematosus: the old and the new. *Clinical & developmental immunology* 2012:823085.

Andrews, S.F., and P.C. Wilson. 2010. The anergic B cell. *Blood* 115:4976-4978.

Andrews, S.F., Q. Zhang, S. Lim, L. Li, J.H. Lee, N.Y. Zheng, M. Huang, W.M. Taylor, A.D. Farris, D. Ni, W. Meng, E.T. Luning Prak, and P.C. Wilson. 2013. Global analysis of

B cell selection using an immunoglobulin light chain-mediated model of autoreactivity. *J Exp Med* 210:125-142.

Anolik, J.H., J.W. Friedberg, B. Zheng, J. Barnard, T. Owen, E. Cushing, J. Kelly, E.C. Milner, R.I. Fisher, and I. Sanz. 2007. B cell reconstitution after rituximab treatment of lymphoma recapitulates B cell ontogeny. *Clin Immunol* 122:139-145.

Arbuckle, M.R., J.A. James, K.F. Kohlhase, M.V. Rubertone, G.J. Dennis, and J.B. Harley. 2001. Development of anti-dsDNA autoantibodies prior to clinical diagnosis of systemic lupus erythematosus. *Scandinavian journal of immunology* 54:211-219.

Aron-Maor, A., and Y. Shoenfeld. 2001. Vaccination and systemic lupus erythematosus: the bidirectional dilemmas. *Lupus* 10:237-240.

Bach, J.F. 2002. The effect of infections on susceptibility to autoimmune and allergic diseases. *The New England journal of medicine* 347:911-920.

Baechler, E.C., P.K. Gregersen, and T.W. Behrens. 2004. The emerging role of interferon in human systemic lupus erythematosus. *Current opinion in immunology* 16:801-807.

Bassing, C.H., F.W. Alt, M.M. Hughes, M. D'Auteuil, T.D. Wehrly, B.B. Woodman, F. Gartner, J.M. White, L. Davidson, and B.P. Sleckman. 2000. Recombination signal sequences restrict chromosomal V(D)J recombination beyond the 12/23 rule. *Nature* 405:583-586.

Basso, K., and R. Dalla-Favera. 2012. Roles of BCL6 in normal and transformed germinal center B cells. *Immunological reviews* 247:172-183.

Basso, K., C. Schneider, Q. Shen, A.B. Holmes, M. Setty, C. Leslie, and R. Dalla-Favera. 2012. BCL6 positively regulates AID and germinal center gene expression via repression of miR-155. *J Exp Med* 209:2455-2465.

Begovich, A.B., V.E. Carlton, L.A. Honigberg, S.J. Schrodi, A.P. Chokkalingam, H.C. Alexander, K.G. Ardlie, Q. Huang, A.M. Smith, J.M. Spoerke, M.T. Conn, M. Chang, S.Y. Chang, R.K. Saiki, J.J. Catanese, D.U. Leong, V.E. Garcia, L.B. McAllister, D.A. Jeffery, A.T. Lee, F. Batliwalla, E. Remmers, L.A. Criswell, M.F. Seldin, D.L. Kastner, C.I. Amos, J.J. Sninsky, and P.K. Gregersen. 2004. A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. *American journal of human genetics* 75:330-337.

Beishuizen, A., M.A. Verhoeven, E.J. Mol, and J.J. van Dongen. 1994. Detection of immunoglobulin kappa light-chain gene rearrangement patterns by Southern blot analysis. *Leukemia* 8:2228-2236; discussion 2237-2229.

Bemark, M., J. Holmqvist, J. Abrahamsson, and K. Mellgren. 2012. Translational Mini-Review Series on B cell subsets in disease. Reconstitution after haematopoietic stem cell transplantation - revelation of B cell developmental pathways and lineage phenotypes. *Clinical and experimental immunology* 167:15-25.

Bentolila, L.A., S. Olson, A. Marshall, F. Rougeon, C.J. Paige, N. Doyen, and G.E. Wu. 1999. Extensive junctional diversity in Ig light chain genes from early B cell progenitors of mu MT mice. *J Immunol* 162:2123-2128.

Berg, J., M. McDowell, H.M. Jack, and M. Wabl. 1990. Immunoglobulin lambda gene rearrangement can precede kappa gene rearrangement. *Developmental immunology* 1:53-57.

Bergman, Y. 1999. Allelic exclusion in B and T lymphopoiesis. *Seminars in immunology* 11:319-328.

Betz, A.G., C. Milstein, A. Gonzalez-Fernandez, R. Pannell, T. Larson, and M.S. Neuberger. 1994. Elements regulating somatic hypermutation of an immunoglobulin kappa

gene: critical role for the intron enhancer/matrix attachment region. *Cell* 77:239-248.

Blair, P.A., K.A. Chavez-Rueda, J.G. Evans, M.J. Shlomchik, A. Eddaoudi, D.A. Isenberg, M.R. Ehrenstein, and C. Mauri. 2009. Selective targeting of B cells with agonistic anti-CD40 is an efficacious strategy for the generation of induced regulatory T2-like B cells and for the suppression of lupus in MRL/lpr mice. *J Immunol* 182:3492-3502.

Blair, P.A., L.Y. Norena, F. Flores-Borja, D.J. Rawlings, D.A. Isenberg, M.R. Ehrenstein, and C. Mauri. 2010. CD19(+)CD24(hi)CD38(hi) B cells exhibit regulatory capacity in healthy individuals but are functionally impaired in systemic Lupus Erythematosus patients. *Immunity* 32:129-140.

Blanchard-Rohner, G., A.S. Pulickal, C.M. Jol-van der Zijde, M.D. Snape, and A.J. Pollard. 2009. Appearance of peripheral blood plasma cells and memory B cells in a primary and secondary immune response in humans. *Blood* 114:4998-5002.

Blanco, P., A.K. Palucka, M. Gill, V. Pascual, and J. Banchereau. 2001. Induction of dendritic cell differentiation by IFN-alpha in systemic lupus erythematosus. *Science* 294:1540-1543.

Bolland, D.J., A.L. Wood, C.M. Johnston, S.F. Bunting, G. Morgan, L. Chakalova, P.J. Fraser, and A.E. Corcoran. 2004. Antisense intergenic transcription in V(D)J recombination. *Nature immunology* 5:630-637.

Bonelli, M., A. Savitskaya, K. von Dalwigk, C.W. Steiner, D. Aletaha, J.S. Smolen, and C. Scheinecker. 2008. Quantitative and qualitative deficiencies of regulatory T cells in patients with systemic lupus erythematosus (SLE). *International immunology* 20:861-868.

Boone, D.L., E.E. Turer, E.G. Lee, R.C. Ahmad, M.T. Wheeler, C. Tsui, P. Hurley, M. Chien, S. Chai, O. Hitotsumatsu, E. McNally, C. Pickart, and A. Ma. 2004. The ubiquitin-modifying enzyme A20 is required for termination of Toll-like receptor responses. *Nature immunology* 5:1052-1060.

Bottini, N., L. Musumeci, A. Alonso, S. Rahmouni, K. Nika, M. Rostamkhani, J. MacMurray, G.F. Meloni, P. Lucarelli, M. Pellecchia, G.S. Eisenbarth, D. Comings, and T. Mustelin. 2004. A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. *Nature genetics* 36:337-338.

Bouaziz, J.D., S. Calbo, M. Maho-Vaillant, A. Saussine, M. Bagot, A. Bensussan, and P. Musette. 2010. IL-10 produced by activated human B cells regulates CD4(+) T-cell activation in vitro. *Eur J Immunol* 40:2686-2691.

Braun, D., I. Caramalho, and J. Demengeot. 2002. IFN-alpha/beta enhances BCR-dependent B cell responses. *International immunology* 14:411-419.

Bräuninger, A., T. Goossens, K. Rajewsky, and R. Küppers. 2001. Regulation of immunoglobulin light chain gene rearrangements during early B cell development in the human. *European Journal of Immunology* 31:3631-3637.

Budarf, M.L., P. Goyette, G. Boucher, J. Lian, R.R. Graham, J.O. Claudio, T. Hudson, D. Gladman, A.E. Clarke, J.E. Pope, C. Peschken, C.D. Smith, J. Hanly, E. Rich, G. Boire, S.G. Barr, M. Zummer, P.R. Fortin, J. Wither, and J.D. Rioux. 2011. A targeted association study in systemic lupus erythematosus identifies multiple susceptibility alleles. *Genes and immunity* 12:51-58.

Cai, X.Y., M. Luo, X.J. Lin, Y.L. Xu, C. Tang, J.H. Ye, W.N. Li, Z.X. He, and N. Yu. 2012. [Expression and significance of Th17 and Treg cells in peripheral blood of patients with systemic lupus erythematosus]. *Zhonghua yi xue za zhi* 92:460-463.

Caraux, A., B. Klein, B. Paiva, C. Bret, A. Schmitz, G.M. Fuhler, N.A. Bos, H.E. Johnsen, A. Orfao, and M. Perez-Andres. 2010. Circulating human B and plasma cells. Age-associated changes in counts and detailed characterization of circulating normal CD138- and CD138+ plasma cells. *Haematologica* 95:1016-1020.

Carey, J.B., C.S. Moffatt-Blue, L.C. Watson, A.L. Gavin, and A.J. Feeney. 2008. Repertoire-based selection into the marginal zone compartment during B cell development. *J Exp Med* 205:2043-2052.

Carter, N.A., E.C. Rosser, and C. Mauri. 2012. Interleukin-10 produced by B cells is crucial for the suppression of Th17/Th1 responses, induction of T regulatory type 1 cells and reduction of collagen-induced arthritis. *Arthritis research & therapy* 14:R32.

Casellas, R., T.A. Shih, M. Kleinewietfeld, J. Rakonjac, D. Nemazee, K. Rajewsky, and M.C. Nussenzweig. 2001. Contribution of receptor editing to the antibody repertoire. *Science* 291:1541-1544.

Casellas, R., Q. Zhang, N.Y. Zheng, M.D. Mathias, K. Smith, and P.C. Wilson. 2007. Igkappa allelic inclusion is a consequence of receptor editing. *J Exp Med* 204:153-160.

Castillejo-Lopez, C., A.M. Delgado-Vega, J. Wojcik, S.V. Kozyrev, E. Thavathiru, Y.Y. Wu, E. Sanchez, D. Pollmann, J.R. Lopez-Egido, S. Fineschi, N. Dominguez, R. Lu, J.A. James, J.T. Merrill, J.A. Kelly, K.M. Kaufman, K.L. Moser, G. Gilkeson, J. Frostegard, B.A. Pons-Estel, S. D'Alfonso, T. Witte, J.L. Callejas, J.B. Harley, P.M. Gaffney, J. Martin, J.M. Guthridge, and M.E. Alarcon-Riquelme. 2012. Genetic and physical interaction of the B-cell systemic lupus erythematosus-associated genes BANK1 and BLK. *Annals of the rheumatic diseases* 71:136-142.

Chailyan, A., P. Marcatili, and A. Tramontano. 2011. The association of heavy and light chain variable domains in antibodies: implications for antigen specificity. *The FEBS journal* 278:2858-2866.

Chan, T.D., D. Gatto, K. Wood, T. Camidge, A. Basten, and R. Brink. 2009. Antigen affinity controls rapid T-dependent antibody production by driving the expansion rather than the differentiation or extrafollicular migration of early plasmablasts. *J Immunol* 183:3139-3149.

Chen, C., M.P. Stenzel-Poore, and M.B. Rittenberg. 1991. Natural auto- and polyreactive antibodies differing from antigen-induced antibodies in the H chain CDR3. *J Immunol* 147:2359-2367.

Cherry, S.R., C. Beard, R. Jaenisch, and D. Baltimore. 2000. V(D)J recombination is not activated by demethylation of the kappa locus. *Proceedings of the National Academy of Sciences of the United States of America* 97:8467-8472.

Chitrabamrung, S., R.L. Rubin, and E.M. Tan. 1981. Serum deoxyribonuclease I and clinical activity in systemic lupus erythematosus. *Rheumatology international* 1:55-60.

Chu, V.T., P. Enghard, S. Schurer, G. Steinhauser, B. Rudolph, G. Riemekasten, and C. Berek. 2009. Systemic activation of the immune system induces aberrant BAFF and APRIL expression in B cells in patients with systemic lupus erythematosus. *Arthritis and rheumatism* 60:2083-2093.

Chung, S.A., and L.A. Criswell. 2007. PTPN22: its role in SLE and autoimmunity. *Autoimmunity* 40:582-590.

Coleclough, C., R.P. Perry, K. Karjalainen, and M. Weigert. 1981. Aberrant rearrangements contribute significantly to the allelic exclusion of immunoglobulin gene expression. *Nature* 290:372-378.

Cooper, G.S., C.G. Parks, E.L. Treadwell, E.W. St Clair, G.S. Gilkeson, P.L. Cohen, R.A. Roubey, and M.A. Dooley. 2002. Differences by race, sex and age in the clinical

and immunologic features of recently diagnosed systemic lupus erythematosus patients in the southeastern United States. *Lupus* 11:161-167.
Corcoran, A.E. 2005. Immunoglobulin locus silencing and allelic exclusion. *Seminars in immunology* 17:141-154.
Corcoran, A.E., A. Riddell, D. Krooshoop, and A.R. Venkitaraman. 1998. Impaired immunoglobulin gene rearrangement in mice lacking the IL-7 receptor. *Nature* 391:904-907.
Corcoran, A.E., F.M. Smart, R.J. Cowling, T. Crompton, M.J. Owen, and A.R. Venkitaraman. 1996. The interleukin-7 receptor alpha chain transmits distinct signals for proliferation and differentiation during B lymphopoiesis. *The EMBO journal* 15:1924-1932.
Crispin, J.C., V.C. Kyttaris, C. Terhorst, and G.C. Tsokos. 2010. T cells as therapeutic targets in SLE. *Nature reviews. Rheumatology* 6:317-325.
Crispin, J.C., M. Oukka, G. Bayliss, R.A. Cohen, C.A. Van Beek, I.E. Stillman, V.C. Kyttaris, Y.T. Juang, and G.C. Tsokos. 2008. Expanded double negative T cells in patients with systemic lupus erythematosus produce IL-17 and infiltrate the kidneys. *J Immunol* 181:8761-8766.
Criswell, L.A. 2008. The genetic contribution to systemic lupus erythematosus. *Bulletin of the NYU hospital for joint diseases* 66:176-183.
Crouzier, R., T. Martin, and J.L. Pasquali. 1995. Heavy chain variable region, light chain variable region, and heavy chain CDR3 influences on the mono- and polyreactivity and on the affinity of human monoclonal rheumatoid factors. *J Immunol* 154:4526-4535.
Crow, M.K., and K.A. Kirou. 2004. Interferon-alpha in systemic lupus erythematosus. *Current opinion in rheumatology* 16:541-547.
Cui, Y., Y. Sheng, and X. Zhang. 2013. Genetic susceptibility to SLE: recent progress from GWAS. *Journal of autoimmunity* 41:25-33.
Culton, D.A., B.P. O'Conner, K.L. Conway, R. Diz, J. Rutan, B.J. Vilen, and S.H. Clarke. 2006. Early preplasma cells define a tolerance checkpoint for autoreactive B cells. *J Immunol* 176:790-802.
Cunninghame Graham, D.S., H. Manku, S. Wagner, J. Reid, K. Timms, A. Gutin, J.S. Lanchbury, and T.J. Vyse. 2007. Association of IRF5 in UK SLE families identifies a variant involved in polyadenylation. *Human molecular genetics* 16:579-591.
D'Cruz, D.P., M.A. Khamashta, and G.R. Hughes. 2007. Systemic lupus erythematosus. *Lancet* 369:587-596.
Daitch, L.E., M.W. Moore, D.M. Persiani, J.M. Durdik, and E. Selsing. 1992. Transcription and recombination of the murine RS element. *J Immunol* 149:832-840.
Das, S., N. Nikolaidis, and M. Nei. 2009. Genomic organization and evolution of immunoglobulin kappa gene enhancers and kappa deleting element in mammals. *Molecular Immunology* 46:3171-3177.
Davids, M.S., M.R. Murali, and D.J. Kuter. 2010. Serum free light chain analysis. *American journal of hematology* 85:787-790.
Deapen, D., A. Escalante, L. Weinrib, D. Horwitz, B. Bachman, P. Roy-Burman, A. Walker, and T.M. Mack. 1992. A revised estimate of twin concordance in systemic lupus erythematosus. *Arthritis and rheumatism* 35:311-318.
Dekosky, B.J., G.C. Ippolito, R.P. Deschner, J.J. Lavinder, Y. Wine, B.M. Rawlings, N. Varadarajan, C. Giesecke, T. Dorner, S.F. Andrews, P.C. Wilson, S.P. Hunicke-Smith, C.G. Willson, A.D. Ellington, and G. Georgiou. 2013. High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. *Nature biotechnology* 31:166-169.

Di Noia, J.M., and M.S. Neuberger. 2007. Molecular mechanisms of antibody somatic hypermutation. *Annual review of biochemistry* 76:1-22.

Dintzis, H.M., R.Z. Dintzis, and B. Vogelstein. 1976. Molecular determinants of immunogenicity: the immunon model of immune response. *Proceedings of the National Academy of Sciences of the United States of America* 73:3671-3675.

Dintzis, R.Z., M.H. Middleton, and H.M. Dintzis. 1983. Studies on the immunogenicity and tolerogenicity of T-independent antigens. *J Immunol* 131:2196-2203.

Dorner, T., S.J. Foster, N.L. Farner, and P.E. Lipsky. 1998. Immunoglobulin kappa chain receptor editing in systemic lupus erythematosus. *The Journal of clinical investigation* 102:688-694.

Dorner, T., C. Giesecke, and P.E. Lipsky. 2011a. Mechanisms of B cell autoimmunity in SLE. *Arthritis research & therapy* 13:243.

Dörner, T., C. Heimbächer, N.L. Farner, and P.E. Lipsky. 1999. Enhanced Mutational Activity of Vκ Gene Rearrangements in Systemic Lupus Erythematosus. *Clinical Immunology* 92:188-196.

Dorner, T., A.M. Jacobi, J. Lee, and P.E. Lipsky. 2011b. Abnormalities of B cell subsets in patients with systemic lupus erythematosus. *Journal of immunological methods* 363:187-197.

Dorner, T., S. Kaschner, A. Hansen, A. Pruss, and P.E. Lipsky. 2001. Perturbations in the impact of mutational activity on Vlambda genes in systemic lupus erythematosus. *Arthritis research* 3:368-374.

Dunn-Walters, D., L. Boursier, and J. Spencer. 2000. Effect of somatic hypermutation on potential N-glycosylation sites in human immunoglobulin heavy chain variable regions. *Mol Immunol* 37:107-113.

Dunn-Walters, D.K., P.G. Isaacson, and J. Spencer. 1995. Analysis of mutations in immunoglobulin heavy chain variable region genes of microdissected marginal zone (MGZ) B cells suggests that the MGZ of human spleen is a reservoir of memory B cells. *J Exp Med* 182:559-566.

Dunnick, W.A., J.T. Collins, J. Shi, G. Westfield, C. Fontaine, P. Hakimpour, and F.N. Papavasiliou. 2009. Switch recombination and somatic hypermutation are controlled by the heavy chain 3' enhancer region. *J Exp Med* 206:2613-2623.

Durdik, J., M.W. Moore, and E. Selsing. 1984. Novel kappa light-chain gene rearrangements in mouse lambda light chain-producing B lymphocytes. *Nature* 307:749-752.

Duty, J.A., P. Szodoray, N.Y. Zheng, K.A. Koelsch, Q. Zhang, M. Swiatkowski, M. Mathias, L. Garman, C. Helms, B. Nakken, K. Smith, A.D. Farris, and P.C. Wilson. 2009. Functional anergy in a subpopulation of naive B cells from healthy humans that express autoreactive immunoglobulin receptors. *J Exp Med* 206:139-151.

Eastman, Q.M., T.M. Leu, and D.G. Schatz. 1996. Initiation of V(D)J recombination in vitro obeying the 12/23 rule. *Nature* 380:85-88.

Ehlich, A., S. Schaal, H. Gu, D. Kitamura, W. Muller, and K. Rajewsky. 1993. Immunoglobulin heavy and light chain genes rearrange independently at early stages of B cell development. *Cell* 72:695-704.

El Shikh, M.E., R.M. El Sayed, A.K. Szakal, and J.G. Tew. 2009. T-independent antibody responses to T-dependent antigens: a novel follicular dendritic cell-dependent activity. *J Immunol* 182:3482-3491.

Elgueta, R., M.J. Benson, V.C. de Vries, A. Wasiuk, Y. Guo, and R.J. Noelle. 2009. Molecular mechanism and function of CD40/CD40L engagement in the immune system. *Immunological reviews* 229:152-172.

Emlen, W., J. Niebur, and R. Kadera. 1994. Accelerated in vitro apoptosis of lymphocytes from patients with systemic lupus erythematosus. *J Immunol* 152:3685-3692.
Evans, J.G., K.A. Chavez-Rueda, A. Eddaoudi, A. Meyer-Bahlburg, D.J. Rawlings, M.R. Ehrenstein, and C. Mauri. 2007. Novel suppressive function of transitional 2 B cells in experimental arthritis. *J Immunol* 178:7868-7878.
Fagarasan, S., and T. Honjo. 2000. T-Independent immune response: new aspects of B cell biology. *Science* 290:89-92.
Fairhurst, A.M., A.E. Wandstrat, and E.K. Wakeland. 2006. Systemic lupus erythematosus: multiple immunological phenotypes in a complex genetic disease. *Advances in immunology* 92:1-69.
Farner, N.L., T. Dörner, and P.E. Lipsky. 1999. Molecular Mechanisms and Selection Influence the Generation of the Human VλJλ Repertoire. *The Journal of Immunology* 162:2137-2145.
Fecteau, J.F., G. Cote, and S. Neron. 2006. A new memory CD27-IgG+ B cell population in peripheral blood expressing VH genes with low frequency of somatic mutation. *J Immunol* 177:3728-3736.
Felsher, D.W., D.T. Ando, and J. Braun. 1991. Independent rearrangement of Ig lambda genes in tissue culture-derived murine B cell lines. *International immunology* 3:711-718.
Fillatreau, S., C.H. Sweenie, M.J. McGeachy, D. Gray, and S.M. Anderton. 2002. B cells regulate autoimmunity by provision of IL-10. *Nature immunology* 3:944-950.
Fink, K. 2012. Origin and Function of Circulating Plasmablasts during Acute Viral Infections. *Frontiers in immunology* 3:78.
Fleming, H.E., and C.J. Paige. 2001. Pre-B cell receptor signaling mediates selective response to IL-7 at the pro-B to pre-B cell transition via an ERK/MAP kinase-dependent pathway. *Immunity* 15:521-531.
Flores-Borja, F., A. Bosma, D. Ng, V. Reddy, M.R. Ehrenstein, D.A. Isenberg, and C. Mauri. 2013. CD19+CD24hiCD38hi B cells maintain regulatory T cells while limiting TH1 and TH17 differentiation. *Science translational medicine* 5:173ra123.
Foster, S.J., H.P. Brezinschek, R.I. Brezinschek, and P.E. Lipsky. 1997. Molecular mechanisms and selective influences that shape the kappa gene repertoire of IgM+ B cells. *The Journal of clinical investigation* 99:1614-1627.
Fournier, E.M., M.G. Velez, K. Leahy, C.L. Swanson, A.V. Rubtsov, R.M. Torres, and R. Pelanda. 2012. Dual-reactive B cells are autoreactive and highly enriched in the plasmablast and memory B cell subsets of autoimmune mice. *J Exp Med* 209:1797-1812.
Friou, G.J. 1958. Identification of the nuclear component of the interaction of lupus erythematosus globulin and nuclei. *J Immunol* 80:476-481.
Fulton, R., and B. Van Ness. 1993. Kappa immunoglobulin promoters and enhancers display developmentally controlled interactions. *Nucleic acids research* 21:4941-4947.
Garraud, O., G. Borhis, G. Badr, S. Degrelle, B. Pozzetto, F. Cognasse, and Y. Richard. 2012. Revisiting the B-cell compartment in mouse and humans: more than one B-cell subset exists in the marginal zone and beyond. *BMC immunology* 13:63.
Gatto, D., and R. Brink. 2010. The germinal center reaction. *The Journal of allergy and clinical immunology* 126:898-907; quiz 908-899.
Gay, D., T. Saunders, S. Camper, and M. Weigert. 1993. Receptor editing: an approach by autoreactive B cells to escape tolerance. *J Exp Med* 177:999-1008.
Gerdes, T., and M. Wabl. 2004. Autoreactivity and allelic inclusion in a B cell nuclear transfer mouse. *Nature immunology* 5:1282-1287.

Gergely, P., Jr., C. Grossman, B. Niland, F. Puskas, H. Neupane, F. Allam, K. Banki, P.E. Phillips, and A. Perl. 2002. Mitochondrial hyperpolarization and ATP depletion in patients with systemic lupus erythematosus. *Arthritis and rheumatism* 46:175-190.

Ghia, P., E. ten Boekel, A.G. Rolink, and F. Melchers. 1998. B-cell development: a comparison between mouse and man. *Immunology today* 19:480-485.

Giachino, C., E. Padovan, and A. Lanzavecchia. 1995. kappa+lambda+ dual receptor B cells are present in the human peripheral repertoire. *J Exp Med* 181:1245-1250.

Giachino, C., E. Padovan, and A. Lanzavecchia. 1998. Re-expression of RAG-1 and RAG-2 genes and evidence for secondary rearrangements in human germinal center B lymphocytes. *Eur J Immunol* 28:3506-3513.

Girschick, H.J., A.C. Grammer, T. Nanki, M. Mayo, and P.E. Lipsky. 2001. RAG1 and RAG2 expression by B cell subsets from human tonsil and peripheral blood. *J Immunol* 166:377-386.

Gladman, D.D., D. Ibanez, and M.B. Urowitz. 2002. Systemic lupus erythematosus disease activity index 2000. *The Journal of rheumatology* 29:288-291.

Gomez, D., P.A. Correa, L.M. Gomez, J. Cadena, J.F. Molina, and J.M. Anaya. 2004. Th1/Th2 cytokines in patients with systemic lupus erythematosus: is tumor necrosis factor alpha protective? *Seminars in arthritis and rheumatism* 33:404-413.

Goodnow, C.C., J. Crosbie, S. Adelstein, T.B. Lavoie, S.J. Smith-Gill, R.A. Brink, H. Pritchard-Briscoe, J.S. Wotherspoon, R.H. Loblay, K. Raphael, and et al. 1988. Altered immunoglobulin expression and functional silencing of self-reactive B lymphocytes in transgenic mice. *Nature* 334:676-682.

Goodnow, C.C., J. Crosbie, H. Jorgensen, R.A. Brink, and A. Basten. 1989. Induction of self-tolerance in mature peripheral B lymphocytes. *Nature* 342:385-391.

Goodnow, C.C., C.G. Vinuesa, K.L. Randall, F. Mackay, and R. Brink. 2010. Control systems and decision making for antibody production. *Nature immunology* 11:681-688.

Graham, R.R., C. Kyogoku, S. Sigurdsson, I.A. Vlasova, L.R. Davies, E.C. Baechler, R.M. Plenge, T. Koeuth, W.A. Ortmann, G. Hom, J.W. Bauer, C. Gillett, N. Burtt, D.S. Cunninghame Graham, R. Onofrio, M. Petri, I. Gunnarsson, E. Svenungsson, L. Ronnblom, G. Nordmark, P.K. Gregersen, K. Moser, P.M. Gaffney, L.A. Criswell, T.J. Vyse, A.C. Syvanen, P.R. Bohjanen, M.J. Daly, T.W. Behrens, and D. Altshuler. 2007. Three functional variants of IFN regulatory factor 5 (IRF5) define risk and protective haplotypes for human lupus. *Proceedings of the National Academy of Sciences of the United States of America* 104:6758-6763.

Graninger, W.B., P.L. Goldman, C.C. Morton, S.J. O'Brien, and S.J. Korsmeyer. 1988. The kappa-deleting element. Germline and rearranged, duplicated and dispersed forms. *J Exp Med* 167:488-501.

Guo, L., H. Deshmukh, R. Lu, G.S. Vidal, J.A. Kelly, K.M. Kaufman, N. Dominguez, W. Klein, X. Kim-Howard, G.R. Bruner, R.H. Scofield, K.L. Moser, P.M. Gaffney, I.M. Dozmorov, G.S. Gilkeson, E.K. Wakeland, Q.Z. Li, C.D. Langefeld, M.C. Marion, A.H. Williams, J. Divers, G.S. Alarcon, E.E. Brown, R.P. Kimberly, J.C. Edberg, R. Ramsey-Goldman, J.D. Reveille, G. McGwin, Jr., L.M. Vila, M.A. Petri, T.J. Vyse, J.T. Merrill, J.A. James, S.K. Nath, J.B. Harley, and J.M. Guthridge. 2009. Replication of the BANK1 genetic association with systemic lupus erythematosus in a European-derived population. *Genes and immunity* 10:531-538.

Hakkim, A., B.G. Furnrohr, K. Amann, B. Laube, U.A. Abed, V. Brinkmann, M. Herrmann, R.E. Voll, and A. Zychlinsky. 2010. Impairment of neutrophil extracellular trap

degradation is associated with lupus nephritis. *Proceedings of the National Academy of Sciences of the United States of America* 107:9813-9818.
Halverson, R., R.M. Torres, and R. Pelanda. 2004. Receptor editing is the main mechanism of B cell tolerance toward membrane antigens. *Nature immunology* 5:645-650.
Hewitt, S.L., B. Yin, Y. Ji, J. Chaumeil, K. Marszalek, J. Tenthorey, G. Salvagiotto, N. Steinel, L.B. Ramsey, J. Ghysdael, M.A. Farrar, B.P. Sleckman, D.G. Schatz, M. Busslinger, C.H. Bassing, and J.A. Skok. 2009. RAG-1 and ATM coordinate monoallelic recombination and nuclear positioning of immunoglobulin loci. *Nature immunology* 10:655-664.
Hieter, P.A., S.J. Korsmeyer, T.A. Waldmann, and P. Leder. 1981. Human immunoglobulin kappa light-chain genes are deleted or rearranged in lambda-producing B cells. *Nature* 290:368-372.
Hikida, M., Y. Nakayama, Y. Yamashita, Y. Kumazawa, S.I. Nishikawa, and H. Ohmori. 1998. Expression of recombination activating genes in germinal center B cells: involvement of interleukin 7 (IL-7) and the IL-7 receptor. *J Exp Med* 188:365-372.
Hirabayashi, Y., Y. Oka, M. Tada, R. Takahashi, and T. Ishii. 2007. A potential trigger of nephritogenic anti-DNA antibodies in lupus nephritis. *Annals of the New York Academy of Sciences* 1108:92-95.
Hochberg, M.C. 1997. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis and rheumatism* 40:1725.
Holman, H.R., and H.G. Kunkel. 1957. Affinity between the lupus erythematosus serum factor and cell nuclei and nucleoprotein. *Science* 126:162-163.
Hsu, E. 2009. V(D)J recombination: of mice and sharks. *Advances in experimental medicine and biology* 650:166-179.
Huang, H., J.F. Kearney, M.J. Grusby, C. Benoist, and D. Mathis. 2006. Induction of tolerance in arthritogenic B cells with receptors of differing affinity for self-antigen. *Proceedings of the National Academy of Sciences of the United States of America* 103:3734-3739.
Inlay, M., F.W. Alt, D. Baltimore, and Y. Xu. 2002. Essential roles of the kappa light chain intronic enhancer and 3' enhancer in kappa rearrangement and demethylation. *Nature immunology* 3:463-468.
Isakov, N. 1997. Immunoreceptor tyrosine-based activation motif (ITAM), a unique module linking antigen and Fc receptors to their signaling cascades. *Journal of leukocyte biology* 61:6-16.
Isenberg, D.A. 2007. BILAG, SLEDAI, SIS, ECLAM, WAM, SLAM .... thank you MAM. *Lupus* 16:849-851.
Jacob, C.O., J. Zhu, D.L. Armstrong, M. Yan, J. Han, X.J. Zhou, J.A. Thomas, A. Reiff, B.L. Myones, J.O. Ojwang, K.M. Kaufman, M. Klein-Gitelman, D. McCurdy, L. Wagner-Weiner, E. Silverman, J. Ziegler, J.A. Kelly, J.T. Merrill, J.B. Harley, R. Ramsey-Goldman, L.M. Vila, S.C. Bae, T.J. Vyse, G.S. Gilkeson, P.M. Gaffney, K.L. Moser, C.D. Langefeld, R. Zidovetzki, and C. Mohan. 2009. Identification of IRAK1 as a risk gene with critical role in the pathogenesis of systemic lupus erythematosus. *Proceedings of the National Academy of Sciences of the United States of America* 106:6256-6261.
Jacob, N., and W. Stohl. 2011. Cytokine disturbances in systemic lupus erythematosus. *Arthritis research & therapy* 13:228.
Jacobi, A.M., W. Huang, T. Wang, W. Freimuth, I. Sanz, R. Furie, M. Mackay, C. Aranow, B. Diamond, and A. Davidson. 2010. Effect of long-term belimumab treatment

on B cells in systemic lupus erythematosus: extension of a phase II, double-blind, placebo-controlled, dose-ranging study. *Arthritis and rheumatism* 62:201-210.
Jacobi, A.M., M. Odendahl, K. Reiter, A. Bruns, G.R. Burmester, A. Radbruch, G. Valet, P.E. Lipsky, and T. Dorner. 2003. Correlation between circulating CD27high plasma cells and disease activity in patients with systemic lupus erythematosus. *Arthritis and rheumatism* 48:1332-1342.
Jacobi, A.M., K. Reiter, M. Mackay, C. Aranow, F. Hiepe, A. Radbruch, A. Hansen, G.R. Burmester, B. Diamond, P.E. Lipsky, and T. Dorner. 2008. Activated memory B cell subsets correlate with disease activity in systemic lupus erythematosus: delineation by expression of CD27, IgD, and CD95. *Arthritis and rheumatism* 58:1762-1773.
Jacobi, A.M., J. Zhang, M. Mackay, C. Aranow, and B. Diamond. 2009. Phenotypic characterization of autoreactive B cells--checkpoints of B cell tolerance in patients with systemic lupus erythematosus. *PloS one* 4:e5776.
James, J.A., B.R. Neas, K.L. Moser, T. Hall, G.R. Bruner, A.L. Sestak, and J.B. Harley. 2001. Systemic lupus erythematosus in adults is associated with previous Epstein-Barr virus exposure. *Arthritis and rheumatism* 44:1122-1126.
Jego, G., R. Bataille, and C. Pellat-Deceunynck. 2001. Interleukin-6 is a growth factor for nonmalignant human plasmablasts. *Blood* 97:1817-1822.
Johnson, K., T. Hashimshony, C.M. Sawai, J.M. Pongubala, J.A. Skok, I. Aifantis, and H. Singh. 2008. Regulation of immunoglobulin light-chain recombination by the transcription factor IRF-4 and the attenuation of interleukin-7 signaling. *Immunity* 28:335-345.
Jones, J.M., and M. Gellert. 2002. Ordered assembly of the V(D)J synaptic complex ensures accurate recombination. *The EMBO journal* 21:4162-4171.
Jordan, N., P.M. Lutalo, and D.P. D'Cruz. 2013. Novel therapeutic agents in clinical development for systemic lupus erythematosus. *BMC medicine* 11:120.
Jung, D., and F.W. Alt. 2004. Unraveling V(D)J recombination; insights into gene regulation. *Cell* 116:299-311.
Juul, L., L. Hougs, V. Andersen, A. Svejgaard, and T. Barington. 1997. The normally expressed kappa immunoglobulin light chain gene repertoire and somatic mutations studied by single-sided specific polymerase chain reaction (PCR); frequent occurrence of features often assigned to autoimmunity. *Clinical and experimental immunology* 109:194-203.
Kameyama, S., M. Horie, T. Kikuchi, T. Omura, A. Tadokoro, T. Takeuchi, I. Nakase, Y. Sugiura, and S. Futaki. 2007. Acid wash in determining cellular uptake of Fab/cell-permeating peptide conjugates. *Biopolymers* 88:98-107.
Katz, S.I., D. Parker, and J.L. Turk. 1974. B-cell suppression of delayed hypersensitivity reactions. *Nature* 251:550-551.
Kawasaki, K., S. Minoshima, E. Nakato, K. Shibuya, A. Shintani, S. Asakawa, T. Sasaki, H.G. Klobeck, G. Combriato, H.G. Zachau, and N. Shimizu. 2001. Evolutionary dynamics of the human immunoglobulin kappa locus and the germline repertoire of the Vkappa genes. *Eur J Immunol* 31:1017-1028.
Kawasaki, K., S. Minoshima, E. Nakato, K. Shibuya, A. Shintani, J.L. Schmeits, J. Wang, and N. Shimizu. 1997. One-megabase sequence analysis of the human immunoglobulin lambda gene locus. *Genome research* 7:250-261.
Kerekov, N.S., N.M. Mihaylova, I. Grozdev, T.A. Todorov, M. Nikolova, M. Baleva, J. Prechl, A. Erdei, and A.I. Tchorbanov. 2011. Elimination of autoreactive B cells in humanized SCID mouse model of SLE. *Eur J Immunol* 41:3301-3311.

Keyna, U., G.B. Beck-Engeser, J. Jongstra, S.E. Applequist, and H.M. Jack. 1995. Surrogate light chain-dependent selection of Ig heavy chain V regions. *J Immunol* 155:5536-5542.

Kindt, G., Osborne. 2007. Kuby's Immunology. Sara Tenney,

Klein, U., R. Kuppers, and K. Rajewsky. 1997. Evidence for a large compartment of IgM-expressing memory B cells in humans. *Blood* 89:1288-1298.

Klein, U., K. Rajewsky, and R. Kuppers. 1998. Human immunoglobulin (Ig)M+IgD+ peripheral blood B cells expressing the CD27 cell surface antigen carry somatically mutated variable region genes: CD27 as a general marker for somatically mutated (memory) B cells. *J Exp Med* 188:1679-1689.

Klein, U., Y. Tu, G.A. Stolovitzky, J.L. Keller, J. Haddad, Jr., V. Miljkovic, G. Cattoretti, A. Califano, and R. Dalla-Favera. 2003. Gene expression dynamics during germinal center transit in B cells. *Annals of the New York Academy of Sciences* 987:166-172.

Klobeck, H.G., and H.G. Zachau. 1986. The human CK gene segment and the kappa deleting element are closely linked. *Nucleic acids research* 14:4591-4603.

Knight, J.S., C. Carmona-Rivera, and M.J. Kaplan. 2012. Proteins derived from neutrophil extracellular traps may serve as self-antigens and mediate organ damage in autoimmune diseases. *Frontiers in immunology* 3:380.

Kocks, C., and K. Rajewsky. 1989. Stable expression and somatic hypermutation of antibody V regions in B-cell developmental pathways. *Annual review of immunology* 7:537-559.

Koethe, S., L. Zander, S. Koster, A. Annan, A. Ebenfelt, J. Spencer, and M. Bemark. 2011. Pivotal advance: CD45RB glycosylation is specifically regulated during human peripheral B cell differentiation. *Journal of leukocyte biology* 90:5-19.

Komori, T., A. Okada, V. Stewart, and F.W. Alt. 1993. Lack of N regions in antigen receptor variable region genes of TdT-deficient lymphocytes. *Science* 261:1171-1175.

Korsmeyer, S.J. 1992. Chromosomal translocations in lymphoid malignancies reveal novel proto-oncogenes. *Annual review of immunology* 10:785-807.

Korsmeyer, S.J., P.A. Hieter, J.V. Ravetch, D.G. Poplack, T.A. Waldmann, and P. Leder. 1981. Developmental hierarchy of immunoglobulin gene rearrangements in human leukemic pre-B-cells. *Proceedings of the National Academy of Sciences of the United States of America* 78:7096-7100.

Korsmeyer, S.J., P.A. Hieter, S.O. Sharrow, C.K. Goldman, P. Leder, and T.A. Waldmann. 1982. Normal human B cells display ordered light chain gene rearrangements and deletions. *J Exp Med* 156:975-985.

Kozyrev, S.V., A.K. Abelson, J. Wojcik, A. Zaghlool, M.V. Linga Reddy, E. Sanchez, I. Gunnarsson, E. Svenungsson, G. Sturfelt, A. Jonsen, L. Truedsson, B.A. Pons-Estel, T. Witte, S. D'Alfonso, N. Barizzone, M.G. Danieli, C. Gutierrez, A. Suarez, P. Junker, H. Laustrup, M.F. Gonzalez-Escribano, J. Martin, H. Abderrahim, and M.E. Alarcon-Riquelme. 2008. Functional variants in the B-cell gene BANK1 are associated with systemic lupus erythematosus. *Nature genetics* 40:211-216.

Kuppers, R., M. Hajadi, L. Plank, K. Rajewsky, and M.L. Hansmann. 1996. Molecular Ig gene analysis reveals that monocytoid B cell lymphoma is a malignancy of mature B cells carrying somatically mutated V region genes and suggests that rearrangement of the kappa-deleting element (resulting in deletion of the Ig kappa enhancers) abolishes somatic hypermutation in the human. *Eur J Immunol* 26:1794-1800.

Kyogoku, C., C.D. Langefeld, W.A. Ortmann, A. Lee, S. Selby, V.E. Carlton, M. Chang, P. Ramos, E.C. Baechler, F.M. Batliwalla, J. Novitzke, A.H. Williams, C. Gillett, P. Rodine, R.R. Graham, K.G. Ardlie, P.M. Gaffney, K.L. Moser, M. Petri, A.B. Begovich, P.K. Gregersen, and T.W. Behrens. 2004. Genetic association of the R620W polymorphism of protein tyrosine phosphatase PTPN22 with human SLE. *American journal of human genetics* 75:504-507.

Kyte, J., and R.F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *Journal of molecular biology* 157:105-132.

Lachmann, P.J., H.J. Muller-Eberhard, H.G. Kunkel, and F. Paronetto. 1962. The localization of in vivo bound complement in tissue section. *J Exp Med* 115:63-82.

Lampropoulou, V., K. Hoehlig, T. Roch, P. Neves, E. Calderon Gomez, C.H. Sweenie, Y. Hao, A.A. Freitas, U. Steinhoff, S.M. Anderton, and S. Fillatreau. 2008. TLR-activated B cells suppress T cell-mediated autoimmunity. *J Immunol* 180:4763-4773.

Lande, R., D. Ganguly, V. Facchinetti, L. Frasca, C. Conrad, J. Gregorio, S. Meller, G. Chamilos, R. Sebasigari, V. Riccieri, R. Bassett, H. Amuro, S. Fukuhara, T. Ito, Y.J. Liu, and M. Gilliet. 2011. Neutrophils activate plasmacytoid dendritic cells by releasing self-DNA-peptide complexes in systemic lupus erythematosus. *Science translational medicine* 3:73ra19.

Lau, C.S., G. Yin, and M.Y. Mok. 2006. Ethnic and geographical differences in systemic lupus erythematosus: an overview. *Lupus* 15:715-719.

Le Bon, A., C. Thompson, E. Kamphuis, V. Durand, C. Rossmann, U. Kalinke, and D.F. Tough. 2006. Cutting edge: enhancement of antibody responses through direct stimulation of B and T cells by type I IFN. *J Immunol* 176:2074-2078.

Leadbetter, E.A., I.R. Rifkin, A.M. Hohlbaum, B.C. Beaudette, M.J. Shlomchik, and A. Marshak-Rothstein. 2002. Chromatin-IgG complexes activate B cells by dual engagement of IgM and Toll-like receptors. *Nature* 416:603-607.

Lee, E.G., D.L. Boone, S. Chai, S.L. Libby, M. Chien, J.P. Lodolce, and A. Ma. 2000. Failure to regulate TNF-induced NF-kappaB and cell death responses in A20-deficient mice. *Science* 289:2350-2354.

Lefranc, M.P. 2001a. Nomenclature of the human immunoglobulin heavy (IGH) genes. *Experimental and clinical immunogenetics* 18:100-116.

Lefranc, M.P. 2001b. Nomenclature of the human immunoglobulin kappa (IGK) genes. *Experimental and clinical immunogenetics* 18:161-174.

Lefranc, M.P. 2001c. Nomenclature of the human immunoglobulin lambda (IGL) genes. *Experimental and clinical immunogenetics* 18:242-254.

Lesinski, G.B., and M.A. Westerink. 2001. Novel vaccine strategies to T-independent antigens. *Journal of microbiological methods* 47:135-149.

Leslie, D., P. Lipsky, and A.L. Notkins. 2001. Autoantibodies as predictors of disease. *The Journal of clinical investigation* 108:1417-1422.

Lewis, S., N. Rosenberg, F. Alt, and D. Baltimore. 1982. Continuing kappa-gene rearrangement in a cell line transformed by Abelson murine leukemia virus. *Cell* 30:807-816.

Li, Y., H. Li, D. Ni, and M. Weigert. 2002a. Anti-DNA B cells in MRL/lpr mice show altered differentiation and editing pattern. *J Exp Med* 196:1543-1552.

Li, Y., H. Li, and M. Weigert. 2002b. Autoreactive B cells in the marginal zone that express dual receptors. *J Exp Med* 195:181-188.

Li, Y.S., K. Hayakawa, and R.R. Hardy. 1993. The regulated expression of B lineage associated genes during B cell differentiation in bone marrow and fetal liver. *J Exp Med* 178:951-960.

Liang, H.E., L.Y. Hsu, D. Cado, and M.S. Schlissel. 2004. Variegated transcriptional activation of the immunoglobulin kappa locus in pre-b cells contributes to the allelic exclusion of light-chain expression. *Cell* 118:19-29.
Liu, S., M.G. Velez, J. Humann, S. Rowland, F.J. Conrad, R. Halverson, R.M. Torres, and R. Pelanda. 2005. Receptor editing can lead to allelic inclusion and development of B cells that retain antibodies reacting with high avidity autoantigens. *J Immunol* 175:5067-5076.
Liu, Z., and A. Davidson. 2011. BAFF and selection of autoreactive B cells. *Trends in immunology* 32:388-394.
Lu, R., K.L. Medina, D.W. Lancki, and H. Singh. 2003. IRF-4,8 orchestrate the pre-B-to-B transition in lymphocyte development. *Genes & development* 17:1703-1708.
Lutalo, P.M., and D.P. D'Cruz. 2012. Belimumab for the management of systemic lupus erythematosus. *Expert opinion on biological therapy* 12:957-963.
MacLennan, I.C., K.M. Toellner, A.F. Cunningham, K. Serre, D.M. Sze, E. Zuniga, M.C. Cook, and C.G. Vinuesa. 2003. Extrafollicular antibody responses. *Immunological reviews* 194:8-18.
Makdasi, E., and D. Eilat. 2013. L chain allelic inclusion does not increase autoreactivity in lupus-prone New Zealand Black/New Zealand White mice. *J Immunol* 190:1472-1480.
Manheimer-Lory, A.J., G. Zandman-Goddard, A. Davidson, C. Aranow, and B. Diamond. 1997. Lupus-specific antibodies reveal an altered pattern of somatic mutation. *The Journal of clinical investigation* 100:2538-2546.
Marie-Paule Lefranc, G.L. 2002. The Immunoglobulin Factsbook. Academic Press (5 Nov 2002),
Mauri, C., C.Q. Chu, D. Woodrow, L. Mori, and M. Londei. 1997. Treatment of a newly established transgenic model of chronic arthritis with nondepleting anti-CD4 monoclonal antibody. *J Immunol* 159:5032-5041.
Mauri, C., D. Gray, N. Mushtaq, and M. Londei. 2003. Prevention of arthritis by interleukin 10-producing B cells. *J Exp Med* 197:489-501.
Mauri, C., L.T. Mars, and M. Londei. 2000. Therapeutic activity of agonistic monoclonal antibodies against CD40 in a chronic autoimmune inflammatory process. *Nature medicine* 6:673-679.
McMurray, R.W. 2001. Estrogen, prolactin, and autoimmunity: actions and interactions. *International immunopharmacology* 1:995-1008.
Mebius, R.E., and G. Kraal. 2005. Structure and function of the spleen. *Nature reviews. Immunology* 5:606-616.
Medzihradszky, K.F. 2008. Characterization of site-specific N-glycosylation. *Methods Mol Biol* 446:293-316.
Meffre, E., R. Casellas, and M.C. Nussenzweig. 2000a. Antibody regulation of B cell development. *Nature immunology* 1:379-385.
Meffre, E., E. Davis, C. Schiff, C. Cunningham-Rundles, L.B. Ivashkiv, L.M. Staudt, J.W. Young, and M.C. Nussenzweig. 2000b. Circulating human B cells that express surrogate light chains and edited receptors. *Nature immunology* 1:207-213.
Meffre, E., A. Schaefer, H. Wardemann, P. Wilson, E. Davis, and M.C. Nussenzweig. 2004. Surrogate light chain expressing human peripheral B cells produce self-reactive antibodies. *J Exp Med* 199:145-150.
Meffre, E., and H. Wardemann. 2008. B-cell tolerance checkpoints in health and autoimmunity. *Current opinion in immunology* 20:632-638.

Mei, H.E., T. Yoshida, W. Sime, F. Hiepe, K. Thiele, R.A. Manz, A. Radbruch, and T. Dorner. 2009. Blood-borne human plasma cells in steady state are derived from mucosal immune responses. *Blood* 113:2461-2469.

Melchers, F., E. ten Boekel, T. Seidl, X.C. Kong, T. Yamagami, K. Onishi, T. Shimizu, A.G. Rolink, and J. Andersson. 2000. Repertoire selection by pre-B-cell receptors and B-cell receptors, and genetic control of B-cell development from immature to mature B cells. *Immunological reviews* 175:33-46.

Merrill, J.T., C.M. Neuwelt, D.J. Wallace, J.C. Shanahan, K.M. Latinis, J.C. Oates, T.O. Utset, C. Gordon, D.A. Isenberg, H.J. Hsieh, D. Zhang, and P.G. Brunetta. 2010. Efficacy and safety of rituximab in moderately-to-severely active systemic lupus erythematosus: the randomized, double-blind, phase II/III systemic lupus erythematosus evaluation of rituximab trial. *Arthritis and rheumatism* 62:222-233.

Meyer, K.B., and J. Ireland. 1998. PMA/ionomycin induces Ig kappa 3' enhancer activity which is in part mediated by a unique NFAT transcription complex. *Eur J Immunol* 28:1467-1480.

Miyara, M., Z. Amoura, C. Parizot, C. Badoual, K. Dorgham, S. Trad, D. Nochy, P. Debre, J.C. Piette, and G. Gorochov. 2005. Global natural regulatory T cell depletion in active systemic lupus erythematosus. *J Immunol* 175:8392-8400.

Miyashita, T., M.J. McIlraith, A.C. Grammer, Y. Miura, J.F. Attrep, Y. Shimaoka, and P.E. Lipsky. 1997. Bidirectional regulation of human B cell responses by CD40-CD40 ligand interactions. *J Immunol* 158:4620-4633.

Mizoguchi, A., E. Mizoguchi, H. Takedatsu, R.S. Blumberg, and A.K. Bhan. 2002. Chronic intestinal inflammatory condition generates IL-10-producing regulatory B cell subset characterized by CD1d upregulation. *Immunity* 16:219-230.

Molina, V., and Y. Shoenfeld. 2005. Infection, vaccines and other environmental triggers of autoimmunity. *Autoimmunity* 38:235-245.

Mond, J.J., A. Lees, and C.M. Snapper. 1995. T cell-independent antigens type 2. *Annual review of immunology* 13:655-692.

Moser, K.L., J.A. Kelly, C.J. Lessard, and J.B. Harley. 2009. Recent insights into the genetic basis of systemic lupus erythematosus. *Genes and immunity* 10:373-379.

Mostoslavsky, R., N. Singh, A. Kirillov, R. Pelanda, H. Cedar, A. Chess, and Y. Bergman. 1998. Kappa chain monoallelic demethylation and the establishment of allelic exclusion. *Genes & development* 12:1801-1811.

Mostoslavsky, R., N. Singh, T. Tenzen, M. Goldmit, C. Gabay, S. Elizur, P. Qi, B.E. Reubinoff, A. Chess, H. Cedar, and Y. Bergman. 2001. Asynchronous replication and allelic exclusion in the immune system. *Nature* 414:221-225.

Mouquet, H., and M.C. Nussenzweig. 2012. Polyreactive antibodies in adaptive immune responses to viruses. *Cellular and molecular life sciences : CMLS* 69:1435-1445.

Muller, B., and M. Reth. 1988. Ordered activation of the Ig lambda locus in Abelson B cell lines. *J Exp Med* 168:2131-2137.

Munoz, L.E., U.S. Gaipl, S. Franz, A. Sheriff, R.E. Voll, J.R. Kalden, and M. Herrmann. 2005. SLE--a disease of clearance deficiency? *Rheumatology (Oxford)* 44:1101-1107.

Munoz, L.E., K. Lauber, M. Schiller, A.A. Manfredi, and M. Herrmann. 2010. The role of defective clearance of apoptotic cells in systemic autoimmunity. *Nature reviews. Rheumatology* 6:280-289.

Murphy. K, T.P., Walport, M. 2008. Janeway's Immunobiology. Taylor & Francis Group, New York and London: Garland Science.

Nagaoka, H., W. Yu, and M.C. Nussenzweig. 2000. Regulation of RAG expression in developing lymphocytes. *Current opinion in immunology* 12:187-190.

Namjou, B., A.L. Sestak, D.L. Armstrong, R. Zidovetzki, J.A. Kelly, N. Jacob, V. Ciobanu, K.M. Kaufman, J.O. Ojwang, J. Ziegler, F.P. Quismorio, Jr., A. Reiff, B.L. Myones, J.M. Guthridge, S.K. Nath, G.R. Bruner, R. Mehrian-Shai, E. Silverman, M. Klein-Gitelman, D. McCurdy, L. Wagner-Weiner, J.J. Nocton, C. Putterman, S.C. Bae, Y.J. Kim, M. Petri, J.D. Reveille, T.J. Vyse, G.S. Gilkeson, D.L. Kamen, M.E. Alarcon-Riquelme, P.M. Gaffney, K.L. Moser, J.T. Merrill, R.H. Scofield, J.A. James, C.D. Langefeld, J.B. Harley, and C.O. Jacob. 2009. High-density genotyping of STAT4 reveals multiple haplotypic associations with systemic lupus erythematosus in different racial groups. *Arthritis and rheumatism* 60:1085-1095.

Nemazee, D. 2000. Receptor selection in B and T lymphocytes. *Annual review of immunology* 18:19-51.

Nemazee, D., and K. Buerki. 1989. Clonal deletion of autoreactive B lymphocytes in bone marrow chimeras. *Proceedings of the National Academy of Sciences of the United States of America* 86:8039-8043.

Nemazee, D., and M. Weigert. 2000. Revising B cell receptors. *J Exp Med* 191:1813-1817.

Nemazee, D.A., and K. Burki. 1989. Clonal deletion of B lymphocytes in a transgenic mouse bearing anti-MHC class I antibody genes. *Nature* 337:562-566.

Neta, R., and S.B. Salvin. 1974. Specific suppression of delayed hypersensitivity: the possible presence of a suppressor B cell in the regulation of delayed hypersensitivity. *J Immunol* 113:1716-1725.

Neuberger, M.S., R.S. Harris, J. Di Noia, and S.K. Petersen-Mahrt. 2003. Immunity through DNA deamination. *Trends in biochemical sciences* 28:305-312.

Newell, K.A., A. Asare, A.D. Kirk, T.D. Gisler, K. Bourcier, M. Suthanthiran, W.J. Burlingham, W.H. Marks, I. Sanz, R.I. Lechler, M.P. Hernandez-Fuentes, L.A. Turka, and V.L. Seyfert-Margolis. 2010. Identification of a B cell signature associated with renal transplant tolerance in humans. *The Journal of clinical investigation* 120:1836-1847.

Nishikomori, R., T. Usui, C.Y. Wu, A. Morinobu, J.J. O'Shea, and W. Strober. 2002. Activated STAT4 has an essential role in Th1 differentiation and proliferation that is independent of its role in the maintenance of IL-12R beta 2 chain expression and signaling. *J Immunol* 169:4388-4398.

Nossal, G.J. 1958. Antibody production by single cells. *British journal of experimental pathology* 39:544-551.

Nossal, G.J., and B.L. Pike. 1980. Clonal anergy: persistence in tolerant mice of antigen-binding B lymphocytes incapable of responding to antigen or mitogen. *Proceedings of the National Academy of Sciences of the United States of America* 77:1602-1606.

Nussenzweig, M.C., A.C. Shaw, E. Sinn, D.B. Danner, K.L. Holmes, H.C. Morse, 3rd, and P. Leder. 1987. Allelic exclusion in transgenic mice that express the membrane form of immunoglobulin mu. *Science* 236:816-819.

O'Connell, F.P., J.L. Pinkus, and G.S. Pinkus. 2004. CD138 (syndecan-1), a plasma cell marker immunohistochemical profile in hematopoietic and nonhematopoietic neoplasms. *American journal of clinical pathology* 121:254-263.

Odendahl, M., A. Jacobi, A. Hansen, E. Feist, F. Hiepe, G.R. Burmester, P.E. Lipsky, A. Radbruch, and T. Dorner. 2000. Disturbed peripheral B lymphocyte homeostasis in systemic lupus erythematosus. *J Immunol* 165:5970-5979.

Oettinger, M.A., D.G. Schatz, C. Gorka, and D. Baltimore. 1990. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science* 248:1517-1523.

Okada, T., M.J. Miller, I. Parker, M.F. Krummel, M. Neighbors, S.B. Hartley, A. O'Garra, M.D. Cahalan, and J.G. Cyster. 2005. Antigen-engaged B cells undergo chemotaxis toward the T zone and form motile conjugates with helper T cells. *PLoS biology* 3:e150.

Orru, V., S.J. Tsai, B. Rueda, E. Fiorillo, S.M. Stanford, J. Dasgupta, J. Hartiala, L. Zhao, N. Ortego-Centeno, S. D'Alfonso, F.C. Arnett, H. Wu, M.A. Gonzalez-Gay, B.P. Tsao, B. Pons-Estel, M.E. Alarcon-Riquelme, Y. He, Z.Y. Zhang, H. Allayee, X.S. Chen, J. Martin, and N. Bottini. 2009. A loss-of-function variant of PTPN22 is associated with reduced risk of systemic lupus erythematosus. *Human molecular genetics* 18:569-579.

Osmond, D.G., S. Rico-Vargas, H. Valenzona, L. Fauteux, L. Liu, R. Janani, L. Lu, and K. Jacobsen. 1994. Apoptosis and macrophage-mediated cell deletion in the regulation of B lymphopoiesis in mouse bone marrow. *Immunological reviews* 142:209-230.

Palanichamy, A., J. Barnard, B. Zheng, T. Owen, T. Quach, C. Wei, R.J. Looney, I. Sanz, and J.H. Anolik. 2009. Novel human transitional B cell populations revealed by B cell depletion therapy. *J Immunol* 182:5982-5993.

Panigrahi, A.K., N.G. Goodman, R.A. Eisenberg, M.R. Rickels, A. Naji, and E.T. Luning Prak. 2008. RS rearrangement frequency as a marker of receptor editing in lupus and type 1 diabetes. *The Journal of experimental medicine* 205:2985-2994.

Pauza, M.E., J.A. Rehmann, and T.W. LeBien. 1993. Unusual patterns of immunoglobulin gene rearrangement and expression during human B cell ontogeny: human B cells can simultaneously express cell surface kappa and lambda light chains. *The Journal of Experimental Medicine* 178:139-149.

Peled, J.U., F.L. Kuang, M.D. Iglesias-Ussel, S. Roa, S.L. Kalis, M.F. Goodman, and M.D. Scharff. 2008. The biochemistry of somatic hypermutation. *Annual review of immunology* 26:481-511.

Pers, J.O., C. Daridon, V. Devauchelle, S. Jousse, A. Saraux, C. Jamin, and P. Youinou. 2005. BAFF overexpression is associated with autoantibody production in autoimmune diseases. *Annals of the New York Academy of Sciences* 1050:34-39.

Persiani, D.M., J. Durdik, and E. Selsing. 1987. Active lambda and kappa antibody gene rearrangement in Abelson murine leukemia virus-transformed pre-B cell lines. *J Exp Med* 165:1655-1674.

Petri, M., W. Stohl, W. Chatham, W.J. McCune, M. Chevrier, J. Ryel, V. Recta, J. Zhong, and W. Freimuth. 2008. Association of plasma B lymphocyte stimulator levels and disease activity in systemic lupus erythematosus. *Arthritis and rheumatism* 58:2453-2459.

Pickering, M.C., M. Botto, P.R. Taylor, P.J. Lachmann, and M.J. Walport. 2000. Systemic lupus erythematosus, complement deficiency, and apoptosis. *Advances in immunology* 76:227-324.

Pillai, S., and D. Baltimore. 1987. Formation of disulphide-linked mu 2 omega 2 tetramers in pre-B cells by the 18K omega-immunoglobulin light chain. *Nature* 329:172-174.

Prokunina, L., C. Castillejo-Lopez, F. Oberg, I. Gunnarsson, L. Berg, V. Magnusson, A.J. Brookes, D. Tentler, H. Kristjansdottir, G. Grondal, A.I. Bolstad, E. Svenungsson, I. Lundberg, G. Sturfelt, A. Jonssen, L. Truedsson, G. Lima, J. Alcocer-Varela, R. Jonsson, U.B. Gyllensten, J.B. Harley, D. Alarcon-Segovia, K. Steinsson, and M.E. Alarcon-Riquelme. 2002. A regulatory polymorphism in PDCD1 is associated with susceptibility to systemic lupus erythematosus in humans. *Nature genetics* 32:666-669.

Radic, M.Z., J. Erikson, S. Litwin, and M. Weigert. 1993. B lymphocytes may escape tolerance by revising their antigen receptors. *J Exp Med* 177:1165-1173.
Radic, M.Z., and M. Weigert. 1995. Origins of anti-DNA antibodies and their implications for B-cell tolerance. *Annals of the New York Academy of Sciences* 764:384-396.
Ramos-Casals, M., M.J. Soto, M.J. Cuadrado, and M.A. Khamashta. 2009. Rituximab in systemic lupus erythematosus: A systematic review of off-label use in 188 cases. *Lupus* 18:767-776.
Ramsden, D.A., and G.E. Wu. 1991. Mouse kappa light-chain recombination signal sequences mediate recombination more frequently than do those of lambda light chain. *Proceedings of the National Academy of Sciences of the United States of America* 88:10721-10725.
Ratcliffe, M.J. 2006. Antibodies, immunoglobulin genes and the bursa of Fabricius in chicken B cell development. *Developmental and comparative immunology* 30:101-118.
Reif, K., E.H. Ekland, L. Ohl, H. Nakano, M. Lipp, R. Forster, and J.G. Cyster. 2002. Balanced responsiveness to chemoattractants from adjacent zones determines B-cell position. *Nature* 416:94-99.
Remmers, E.F., R.M. Plenge, A.T. Lee, R.R. Graham, G. Hom, T.W. Behrens, P.I. de Bakker, J.M. Le, H.S. Lee, F. Batliwalla, W. Li, S.L. Masters, M.G. Booty, J.P. Carulli, L. Padyukov, L. Alfredsson, L. Klareskog, W.V. Chen, C.I. Amos, L.A. Criswell, M.F. Seldin, D.L. Kastner, and P.K. Gregersen. 2007. STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. *The New England journal of medicine* 357:977-986.
Reth, M. 1992. Antigen receptors on B lymphocytes. *Annual review of immunology* 10:97-121.
Reth, M., E. Petrac, P. Wiese, L. Lobel, and F.W. Alt. 1987. Activation of V kappa gene rearrangement in pre-B cells follows the expression of membrane-bound immunoglobulin heavy chains. *The EMBO journal* 6:3299-3305.
Retter, M.W., and D. Nemazee. 1998. Receptor editing occurs frequently during normal B cell development. *J Exp Med* 188:1231-1238.
Revy, P., T. Muto, Y. Levy, F. Geissmann, A. Plebani, O. Sanal, N. Catalan, M. Forveille, R. Dufourcq-Labelouse, A. Gennery, I. Tezcan, F. Ersoy, H. Kayserili, A.G. Ugazio, N. Brousse, M. Muramatsu, L.D. Notarangelo, K. Kinoshita, T. Honjo, A. Fischer, and A. Durandy. 2000. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell* 102:565-575.
Ronnblom, L., and V. Pascual. 2008. The innate immune system in SLE: type I interferons and dendritic cells. *Lupus* 17:394-399.
Roth, D.B., J.P. Menetski, P.B. Nakajima, M.J. Bosma, and M. Gellert. 1992. V(D)J recombination: broken DNA molecules with covalently sealed (hairpin) coding ends in scid mouse thymocytes. *Cell* 70:983-991.
Roth, D.B., C. Zhu, and M. Gellert. 1993. Characterization of broken DNA molecules associated with V(D)J recombination. *Proceedings of the National Academy of Sciences of the United States of America* 90:10788-10792.
Sakano, H., K. Huppi, G. Heinrich, and S. Tonegawa. 1979. Sequences at the somatic recombination sites of immunoglobulin light-chain genes. *Nature* 280:288-294.
Sakurai, T., T. Fujita, I. Kono, T. Kabashima, K. Yamane, N. Tamura, and H. Kashiwagi. 1982. Complement-mediated solubilization of immune complexes in systemic lupus erythematosus. *Clinical and experimental immunology* 48:37-42.

Santoro, D., G. Vita, R. Vita, A. Mallamace, V. Savica, G. Bellinghieri, S. Benvenga, and S. Gangemi. 2010. HLA haplotype in a patient with systemic lupus erythematosus triggered by hepatitis B vaccine. *Clinical nephrology* 74:150-153.

Schatz, D.G., and Y. Ji. 2011. Recombination centres and the orchestration of V(D)J recombination. *Nature reviews. Immunology* 11:251-263.

Schatz, D.G., M.A. Oettinger, and D. Baltimore. 1989. The V(D)J recombination activating gene, RAG-1. *Cell* 59:1035-1048.

Schatz, D.G., M.A. Oettinger, and M.S. Schlissel. 1992. V(D)J recombination: molecular biology and regulation. *Annual review of immunology* 10:359-383.

Schlissel, M.S., and D. Baltimore. 1989. Activation of immunoglobulin kappa gene rearrangement correlates with induction of germline kappa gene transcription. *Cell* 58:1001-1007.

Schlissel, M.S., L.M. Corcoran, and D. Baltimore. 1991. Virus-transformed pre-B cells show ordered activation but not inactivation of immunoglobulin gene rearrangement and transcription. *J Exp Med* 173:711-720.

Schlissel, M.S., and T. Morrow. 1994. Ig heavy chain protein controls B cell development by regulating germ-line transcription and retargeting V(D)J recombination. *J Immunol* 153:1645-1657.

Seifert, M., and R. Kuppers. 2009. Molecular footprints of a germinal center derivation of human IgM+(IgD+)CD27+ B cells and the dynamics of memory B cell generation. *J Exp Med* 206:2659-2669.

Shaffer, A.L., K.I. Lin, T.C. Kuo, X. Yu, E.M. Hurt, A. Rosenwald, J.M. Giltnane, L. Yang, H. Zhao, K. Calame, and L.M. Staudt. 2002. Blimp-1 orchestrates plasma cell differentiation by extinguishing the mature B cell gene expression program. *Immunity* 17:51-62.

Siminovitch, K.A., A. Bakhshi, P. Goldman, and S.J. Korsmeyer. 1985. A uniform deleting element mediates the loss of kappa genes in human B cells. *Nature* 316:260-262.

Siminovitch, K.A., M.W. Moore, J. Durdik, and E. Selsing. 1987. The human kappa deleting element and the mouse recombining segment share DNA sequence homology. *Nucleic acids research* 15:2699-2705.

Sims, G.P., R. Ettinger, Y. Shirota, C.H. Yarboro, G.G. Illei, and P.E. Lipsky. 2005. Identification and characterization of circulating human transitional B cells. *Blood* 105:4390-4398.

Sonoda, E., Y. Pewzner-Jung, S. Schwers, S. Taki, S. Jung, D. Eilat, and K. Rajewsky. 1997. B cell development under the condition of allelic inclusion. *Immunity* 6:225-233.

Stanhope-Baker, P., K.M. Hudson, A.L. Shaffer, A. Constantinescu, and M.S. Schlissel. 1996. Cell type-specific chromatin structure determines the targeting of V(D)J recombinase activity in vitro. *Cell* 85:887-897.

Stavnezer, J. 2011. Complex regulation and function of activation-induced cytidine deaminase. *Trends in immunology* 32:194-201.

Stavnezer, J., J.E. Guikema, and C.E. Schrader. 2008. Mechanism and regulation of class switch recombination. *Annual review of immunology* 26:261-292.

Steinman, R.M., and M.C. Nussenzweig. 2002. Avoiding horror autotoxicus: the importance of dendritic cells in peripheral T cell tolerance. *Proceedings of the National Academy of Sciences of the United States of America* 99:351-358.

Steinman, R.M., M. Pack, and K. Inaba. 1997. Dendritic cells in the T-cell areas of lymphoid organs. *Immunological reviews* 156:25-37.

Stohl, W., A.S. Hamilton, D.M. Deapen, T.M. Mack, and D.A. Horwitz. 1999. Impaired cytotoxic T lymphocyte activity in systemic lupus erythematosus following in

vitro polyclonal T cell stimulation: a contributory role for non-T cells. *Lupus* 8:293-299.

Storb, U., and B. Arp. 1983. Methylation patterns of immunoglobulin genes in lymphoid cells: correlation of expression and differentiation with undermethylation. *Proceedings of the National Academy of Sciences of the United States of America* 80:6642-6646.

Storb, U., P. Engler, J. Manz, K. Gollahon, K. Denis, D. Lo, and R. Brinster. 1988. Expression of immunoglobulin genes in transgenic mice and transfected cells. *Annals of the New York Academy of Sciences* 546:51-56.

Studnicka-Benke, A., G. Steiner, P. Petera, and J.S. Smolen. 1996. Tumour necrosis factor alpha and its soluble receptors parallel clinical disease and autoimmune activity in systemic lupus erythematosus. *British journal of rheumatology* 35:1067-1074.

Su, W., L. Boursier, A. Padala, J.D. Sanderson, and J. Spencer. 2004. Biases in Ig lambda light chain rearrangements in human intestinal plasma cells. *J Immunol* 172:2360-2366.

Su, W., J.N. Gordon, F. Barone, L. Boursier, W. Turnbull, S. Mendis, D.K. Dunn-Walters, and J. Spencer. 2008. Lambda Light Chain Revision in the Human Intestinal IgA Response. *The Journal of Immunology* 181:1264-1271.

Suryani, S., D.A. Fulcher, B. Santner-Nanan, R. Nanan, M. Wong, P.J. Shaw, J. Gibson, A. Williams, and S.G. Tangye. 2010. Differential expression of CD21 identifies developmentally and functionally distinct subsets of human transitional B cells. *Blood* 115:519-529.

Takeda, S., Y.R. Zou, H. Bluethmann, D. Kitamura, U. Muller, and K. Rajewsky. 1993. Deletion of the immunoglobulin kappa chain intron enhancer abolishes kappa chain gene rearrangement in cis but not lambda chain gene rearrangement in trans. *The EMBO journal* 12:2329-2336.

Takeno, M., H. Nagafuchi, S. Kaneko, S. Wakisaka, K. Oneda, Y. Takeba, N. Yamashita, N. Suzuki, H. Kaneoka, and T. Sakane. 1997. Autoreactive T cell clones from patients with systemic lupus erythematosus support polyclonal autoantibody production. *J Immunol* 158:3529-3538.

Tan, E.M., A.S. Cohen, J.F. Fries, A.T. Masi, D.J. McShane, N.F. Rothfield, J.G. Schaller, N. Talal, and R.J. Winchester. 1982. The 1982 revised criteria for the classification of systemic lupus erythematosus. *Arthritis and rheumatism* 25:1271-1277.

Tang, J.Q., M.C. Bene, and G.C. Faure. 1991. Alternative rearrangements of immunoglobulin light chain genes in human leukemia. *Leukemia* 5:651-656.

Tangye, S.G., D.T. Avery, E.K. Deenick, and P.D. Hodgkin. 2003. Intrinsic differences in the proliferation of naive and memory human B cells as a mechanism for enhanced secondary immune responses. *J Immunol* 170:686-694.

Tao, M.H., and S.L. Morrison. 1989. Studies of aglycosylated chimeric mouse-human IgG. Role of carbohydrate in the structure and effector functions mediated by the human IgG constant region. *J Immunol* 143:2595-2601.

Tavares, R.M., E.E. Turer, C.L. Liu, R. Advincula, P. Scapini, L. Rhee, J. Barrera, C.A. Lowell, P.J. Utz, B.A. Malynn, and A. Ma. 2010. The ubiquitin modifying enzyme A20 restricts B cell survival and prevents autoimmunity. *Immunity* 33:181-191.

Taylor, K.E., E.F. Remmers, A.T. Lee, W.A. Ortmann, R.M. Plenge, C. Tian, S.A. Chung, J. Nititham, G. Hom, A.H. Kao, F.Y. Demirci, M.I. Kamboh, M. Petri, S. Manzi, D.L. Kastner, M.F. Seldin, P.K. Gregersen, T.W. Behrens, and L.A. Criswell. 2008. Specificity of the STAT4 genetic association for severe disease manifestations of systemic lupus erythematosus. *PLoS genetics* 4:e1000084.

Teague, B.N., Y. Pan, P.A. Mudd, B. Nakken, Q. Zhang, P. Szodoray, X. Kim-Howard, P.C. Wilson, and A.D. Farris. 2007. Cutting edge: Transitional T3 B cells do not give rise to mature B cells, have undergone selection, and are reduced in murine lupus. *J Immunol* 178:7511-7515.

ten Boekel, E., F. Melchers, and A.G. Rolink. 1998. Precursor B cells showing H chain allelic inclusion display allelic exclusion at the level of pre-B cell receptor surface expression. *Immunity* 8:199-207.

Tiegs, S.L., D.M. Russell, and D. Nemazee. 1993. Receptor editing in self-reactive bone marrow B cells. *J Exp Med* 177:1009-1020.

Tonegawa, S. 1983. Somatic generation of antibody diversity. *Nature* 302:575-581.

Townsend, M.J., J.G. Monroe, and A.C. Chan. 2010. B-cell targeted therapies in human autoimmune diseases: an updated perspective. *Immunological reviews* 237:264-283.

Tsubata, T., and M. Reth. 1990. The products of pre-B cell-specific genes (lambda 5 and VpreB) and the immunoglobulin mu chain form a complex that is transported onto the cell surface. *J Exp Med* 172:973-976.

Uccellini, M.B., L. Busconi, N.M. Green, P. Busto, S.R. Christensen, M.J. Shlomchik, A. Marshak-Rothstein, and G.A. Viglianti. 2008. Autoreactive B cells discriminate CpG-rich and CpG-poor DNA and this response is modulated by IFN-alpha. *J Immunol* 181:5875-5884.

Vassilopoulos, D., B. Kovacs, and G.C. Tsokos. 1995. TCR/CD3 complex-mediated signal transduction pathway in T cells and T cell lines from patients with systemic lupus erythematosus. *J Immunol* 155:2269-2281.

Vela, J.L., D. Aït-Azzouzene, B.H. Duong, T. Ota, and D. Nemazee. 2008. Rearrangement of Mouse Immunoglobulin Kappa Deleting Element Recombining Sequence Promotes Immune Tolerance and Lambda B Cell Production. *Immunity* 28:161-170.

Velez, M.G., M. Kane, S. Liu, S.B. Gauld, J.C. Cambier, R.M. Torres, and R. Pelanda. 2007. Ig allotypic inclusion does not prevent B cell development or response. *J Immunol* 179:1049-1057.

Vettermann, C., and M.S. Schlissel. 2010. Allelic exclusion of immunoglobulin genes: models and mechanisms. *Immunological reviews* 237:22-42.

Vossenkamper, A., P.M. Lutalo, and J. Spencer. 2012. Translational Mini-Review Series on B cell subsets in disease. Transitional B cells in systemic lupus erythematosus and Sjogren's syndrome: clinical implications and effects of B cell-targeted therapies. *Clinical and experimental immunology* 167:7-14.

Vossenkamper, A.S., J. 2013. A role for gut-associated lymphoid tissue in shaping the human B cell repertoire. *Journal of Experimental Medicine, in press*

Walport, M.J. 2002. Complement and systemic lupus erythematosus. *Arthritis research* 4 Suppl 3:S279-293.

Wardemann, H., S. Yurasov, A. Schaefer, J.W. Young, E. Meffre, and M.C. Nussenzweig. 2003. Predominant autoantibody production by early human B cell precursors. *Science* 301:1374-1377.

Watford, W.T., B.D. Hissong, J.H. Bream, Y. Kanno, L. Muul, and J.J. O'Shea. 2004. Signaling by IL-12 and IL-23 and the immunoregulatory roles of STAT4. *Immunological reviews* 202:139-156.

Wei, C., J. Anolik, A. Cappione, B. Zheng, A. Pugh-Bernard, J. Brooks, E.H. Lee, E.C. Milner, and I. Sanz. 2007. A new population of cells lacking expression of CD27 represents a notable component of the B cell memory compartment in systemic lupus erythematosus. *J Immunol* 178:6624-6633.

Weigert, M., R. Perry, D. Kelley, T. Hunkapiller, J. Schilling, and L. Hood. 1980. The joining of V and J gene segments creates antibody diversity. *Nature* 283:497-499.

Weill, J.C., S. Weller, and C.A. Reynaud. 2009. Human marginal zone B cells. *Annual review of immunology* 27:267-285.

Weiner, G.J. 2010. Rituximab: mechanism of action. *Seminars in hematology* 47:115-123.

West, K.L., N.C. Singha, P. De Ioannes, L. Lacomis, H. Erdjument-Bromage, P. Tempst, and P. Cortes. 2005. A direct interaction between the RAG2 C terminus and the core histones is required for efficient V(D)J recombination. *Immunity* 23:203-212.

Wirths, S., and A. Lanzavecchia. 2005. ABCB1 transporter discriminates human resting naive B cells from cycling transitional and memory B cells. *Eur J Immunol* 35:3433-3441.

Wolf, S.D., B.N. Dittel, F. Hardardottir, and C.A. Janeway, Jr. 1996. Experimental autoimmune encephalomyelitis induction in genetically B cell-deficient mice. *J Exp Med* 184:2271-2278.

Wrammert, J., K. Smith, J. Miller, W.A. Langley, K. Kokko, C. Larsen, N.Y. Zheng, I. Mays, L. Garman, C. Helms, J. James, G.M. Air, J.D. Capra, R. Ahmed, and P.C. Wilson. 2008. Rapid cloning of high-affinity human monoclonal antibodies against influenza virus. *Nature* 453:667-671.

Wu, Y.C., D. Kipling, and D.K. Dunn-Walters. 2011. The relationship between CD27 negative and positive B cell populations in human peripheral blood. *Frontiers in immunology* 2:81.

Wu, Y.C., D. Kipling, H.S. Leong, V. Martin, A.A. Ademokun, and D.K. Dunn-Walters. 2010. High-throughput immunoglobulin repertoire analysis distinguishes between human IgM memory and switched memory B-cell populations. *Blood* 116:1070-1078.

Yadav, M., C. Louvet, D. Davini, J.M. Gardner, M. Martinez-Llordella, S. Bailey-Bucktrout, B.A. Anthony, F.M. Sverdrup, R. Head, D.J. Kuster, P. Ruminski, D. Weiss, D. Von Schack, and J.A. Bluestone. 2012. Neuropilin-1 distinguishes natural and inducible regulatory T cells among regulatory T cell subsets in vivo. *J Exp Med* 209:1713-1722, S1711-1719.

Yancopoulos, G.D., and F.W. Alt. 1985. Developmentally controlled and tissue-specific expression of unrearranged VH gene segments. *Cell* 40:271-281.

Yancopoulos, G.D., and F.W. Alt. 1986. Regulation of the assembly and expression of variable-region genes. *Annual review of immunology* 4:339-368.

Yang, M., K. Rui, S. Wang, and L. Lu. 2013. Regulatory B cells in autoimmune diseases. *Cellular & molecular immunology* 10:122-132.

Yee, C.S., D.A. Isenberg, A. Prabu, K. Sokoll, L.S. Teh, A. Rahman, I.N. Bruce, B. Griffiths, M. Akil, N. McHugh, D. D'Cruz, M.A. Khamashta, P. Maddison, A. Zoma, and C. Gordon. 2008. BILAG-2004 index captures systemic lupus erythematosus disease activity better than SLEDAI-2000. *Annals of the rheumatic diseases* 67:873-876.

Yong, P.F., and D.P. D'Cruz. 2008. Mycophenolate mofetil in the treatment of lupus nephritis. *Biologics : targets & therapy* 2:297-310.

Yurasov, S., T. Tiller, M. Tsuiji, K. Velinzon, V. Pascual, H. Wardemann, and M.C. Nussenzweig. 2006. Persistent expression of autoantibodies in SLE patients in remission. *J Exp Med* 203:2255-2261.

Yurasov, S., H. Wardemann, J. Hammersen, M. Tsuiji, E. Meffre, V. Pascual, and M.C. Nussenzweig. 2005. Defective B cell tolerance checkpoints in systemic lupus erythematosus. *J Exp Med* 201:703-711.

Zhu, D., H. McCarthy, C.H. Ottensmeier, P. Johnson, T.J. Hamblin, and F.K. Stevenson. 2002. Acquisition of potential N-glycosylation sites in the immunoglobulin variable region by somatic mutation is a distinctive feature of follicular lymphoma. *Blood* 99:2562-2568.

Zou, Y.R., S. Takeda, and K. Rajewsky. 1993. Gene targeting in the Ig kappa locus: efficient generation of lambda chain-expressing B cells, independent of gene rearrangements in Ig kappa. *The EMBO journal* 12:811-820.